

Db 649 FPPKPKDTLMISRTPEVTCVVVDVSHEDPEVKFNWYVDGVEVHNAKTKPREEQYNSTYRV 708  
 QY 102 VSVLTVLHQDLNGKEYCKKVSNNKALPAPIEKTISKAKGQPREPQVYTLPPSRDELTKNQ 161  
 Db 709 VSVLTVLHQDLNGKEYCKKVSNNKALPAPIEKTISKAKGQPREPQVYTLPPSRDELTKNQ 768  
 QY 162 VSLTCLVKGYFSPDAVAVESNGQPENNYKTTPPVLDSDGSFFLYSKLTVDKSRWQQGNV 221  
 Db 769 VSLTCLVKGYFSPDAVAVESNGQPENNYKTTPPVLDSDGSFFLYSKLTVDKSRWQQGNV 828  
 QY 222 FSCSVMEALHNHYTQKSLSLSPGK 246  
 Db 829 FSCSVMEALHNHYTQKSLSLSPGK 853

RESULT 14  
 AAY92184  
 ID AAY92184 standard; protein; 859 AA.  
 AC  
 XX AAY92184;  
 DT 01-AUG-2000 (first entry)  
 XX Human gp130-Fc-His6.  
 XX  
 KW gp130-Fc-His6; cytokine; antagonist; CNTF; receptor; fusion protein;  
 KW cytostatic; immunomodulator; osteopathic.  
 XX  
 OS Synthetic.  
 OS Homo sapiens.

Key Location/Qualifiers  
 FH Protein 1..619  
 FT /label= gp130  
 FT Peptide 1..22  
 FT /label= signal\_peptide  
 FT /note= "Ser-Gly bridge"  
 FT Disulfide-bond 632  
 FT /note= "forms inter-chain disulfide bridge that link two FC domains"  
 FT Disulfide-bond 635  
 FT /note= "forms inter-chain disulfide bridge that link two FC domains"  
 FT Protein 662..853  
 FT /label= IgG1\_Fc\_domain  
 FT Peptide 854..859  
 FT /label= histidine\_tag

WC200018932-A2.  
 XX  
 XX  
 XX 06-APR-2000.  
 XX  
 XX 22-SEP-1999; 99NO-US22045.  
 XX  
 XX 25-SEP-1998; 98US-0101858.  
 XX 19-MAY-1999; 99US-0313942.  
 XX  
 XX (REGE-) REGENERON PHARM INC.  
 XX  
 XX Stahl N, Yancopoulos GD;  
 XX  
 XX WPI; 2000-293165/25.  
 XX  
 XX Isolated nucleic acid molecule for treating cytokine-related diseases  
 PT or disorders encodes a fusion polypeptide capable of binding a cytokine  
 PT to form a nonfunctional complex  
 XX  
 XX Example 3; Fig 4; 152pp; English.  
 XX  
 XX The invention concerns production of antagonists to any cytokine that  
 CC utilizes an alpha specificity determining component, which when combined  
 CC with the cytokine, binds to a first beta signal transducing component to

CC form a non-functional intermediate which then binds to a second beta  
 CC signal transducing component causing beta-receptor dimerization, the  
 CC soluble alpha specificity determining component of the receptor  
 CC (SR-alpha) and the extracellular domain of the first beta signal  
 CC transducing component of the cytokine receptor (beta-1) are combined to  
 CC form heterodimers (SR-alpha:beta-1) that act as antagonist to the  
 CC cytokine by binding the cytokine to form a non-functional complex. The  
 CC receptor components are shared by cytokines such as the CNTF (ciliary  
 CC neurotrophic factor) family of cytokines. The invention provides the  
 CC basis for the development of IL-6 antagonists, as they show that if, in  
 CC the presence of a ligand, a non-functional intermediate complex,  
 CC consisting of the ligand, its alpha receptor and its beta-1 receptor  
 CC component, can be formed, it will effectively block the action of the  
 CC ligand. Effective antagonists of IL-6 or CNTF consist of heterodimers  
 CC of the extracellular domains of the alpha specificity determining  
 CC components of their receptors and the extracellular domain of gp130.  
 CC The resultant heterodimers, function as high-affinity traps, rendering  
 CC the cytokine inaccessible to form a signal transducing complex with the  
 CC native membrane-bound forms of their receptor. The nucleic acids and  
 CC polypeptides are useful for treating cytokine-related diseases or  
 CC disorders such as osteoporosis and primary and secondary effects of  
 CC cancer including multiple myeloma or cachexia.  
 XX  
 SQ Sequence 859 AA;  
 Query Match 94.0%; Score 1254; DB 21; Length 859;  
 Best Local Similarity 60.5%; Pred. No. 1e-35;  
 Matches 233; Conservative 8; Mismatches 5; Indels 139; Gaps 3;  
 QY 1 DW-----LKAF----- 6  
 Db 469 DWQEDGTVHRTVLRGNLAESKCYLITVPYADGPGSPESIKAYLKQAPPSKGTVRTK 528  
 QY 7 -----YDKVAEKLKEAFM----- 19  
 Db 529 KVGKNEAVLEWDQLPVDVQNGFIRNYTIFYRTIIGNETAVNVDSHTEYTLSSLTSDTL 588  
 QY 20 -----DKHTGCPGPAPELLGGPSVFL 41  
 Db 589 MVRMAAYTDGGKDGPEFTTTPKFAQGEIESGEPKSCDKTHCPGPAPELLGGPSVFL 648  
 QY 42 FPPKPKDTLMISRTPEVTCVVVDVSHEDPEVKFNWYVDGVEVHNAKTKPREEQYNSTYRV 101  
 Db 649 FPPKPKDTLMISRTPEVTCVVVDVSHEDPEVKFNWYVDGVEVHNAKTKPREEQYNSTYRV 708  
 QY 102 VSVLTVLHQDLNGKEYCKKVSNNKALPAPIEKTISKAKGQPREPQVYTLPPSRDELTKNQ 161  
 Db 709 VSVLTVLHQDLNGKEYCKKVSNNKALPAPIEKTISKAKGQPREPQVYTLPPSRDELTKNQ 768  
 QY 162 VSLTCLVKGYFSPDAVAVESNGQPENNYKTTPPVLDSDGSFFLYSKLTVDKSRWQQGNV 221  
 Db 769 VSLTCLVKGYFSPDAVAVESNGQPENNYKTTPPVLDSDGSFFLYSKLTVDKSRWQQGNV 828  
 QY 222 FSCSVMEALHNHYTQKSLSLSPGK 246  
 Db 829 FSCSVMEALHNHYTQKSLSLSPGK 853

RESULT 15  
 AAY92205  
 ID AAY92205 standard; Protein; 1158 AA.  
 XX  
 AC AAY92205;  
 XX  
 DT 01-AUG-2000 (first entry)  
 XX  
 XX Fusion polypeptide 616, IL-6 trap.  
 XX  
 KW IL-6 trap; cytokine; antagonist; CNTF; receptor; fusion protein;  
 KW cytostatic; immunomodulator; osteopathic.  
 XX  
 OS Synthetic.  
 OS Homo sapiens.

XX PN WO200018932-A2.  
XX PD 06-APR-2000.  
XX PF 22-SEP-1999; 99WO-US22045.  
XX PR 25-SEP-1998; 98US-0101858.  
XX PR 19-MAY-1999; 98US-0313942.  
XX PA (REGE-) REGENERON PHARM INC.  
XX PI Stahl N, Yancopoulos GD;  
XX DR WPI; 2000-293165/25.  
XX DR N-PSDB; AAA09047.  
XX PT Isolated nucleic acid molecule for treating cytokine-related diseases  
XX PT or disorders encodes a fusion polypeptide capable of binding a cytokine  
XX PT to form a nonfunctional complex  
XX PS Example 6; Fig 25A-F; 152pp; English.  
XX CC This sequence shows fusion polypeptide 616, which is capable of  
XX CC binding cytokine IL-6 to form a non-functional complex.  
XX CC The invention concerns production of antagonists to any cytokine that  
XX CC utilizes an alpha specificity determining component, which when combined  
XX CC with the cytokine, binds to a first beta signal transducing component to  
XX CC form a non-functional intermediate which then binds to a second beta  
XX CC signal transducing component causing beta-receptor dimerization, the  
XX CC soluble alpha specificity determining component of the receptor  
XX CC (SR-alpha) and the extracellular domain of the first beta signal  
XX CC transducing component of the cytokine receptor (beta-1) are combined to  
XX CC form heterodimers (SR-alpha:beta-1) that act as antagonist to the  
XX CC cytokine by binding the cytokine to form a non-functional complex. The  
XX CC receptor components are shared by cytokines such as the CNRF (ciliary  
XX CC neurotrophic factor) family of cytokines. The invention provides the  
XX CC basis for the development of IL-6 antagonists, as they show that if, in  
XX CC the presence of a ligand, a non-functional intermediate complex,  
XX CC consisting of the ligand, its alpha receptor and its beta-1 receptor  
XX CC component, can be formed, it will effectively block the action of the  
XX CC ligand. Effective antagonists of IL-6 or CNRF consist of heterodimers  
XX CC of the extracellular domains of the alpha specificity determining  
XX CC components of their receptors and the extracellular domain of gp130.  
XX CC The resultant heterodimers, function as high-affinity traps, rendering  
XX CC the cytokine inaccessible to form a signal transducing complex with the  
XX CC native membrane-bound forms of their receptor. The nucleic acids and  
XX CC polypeptides are useful for treating cytokine-related diseases or  
XX CC disorders such as osteoporosis and primary and secondary effects of  
XX CC cancer including multiple myeloma or cachexia.  
XX SQ Sequence 1158 AA;

Query Match 94.0%; Score 1254; DB 21; Length 1158;  
Best Local Similarity 61.2%; Pred. No. 2e-35;  
Matches 233; Conservative 8; Mismatches 5; Indels 135; Gaps 3;

QY 1 DW-----LKA-----6  
DB 778 DMQOEDGTVHRTVLRNLAESKYLITVTPVYADGPGSPESIKAYLKQAPPSKGTVRTK 837  
QY 7 -----YDKVAEKLKEAF-----19  
DB 838 KVGKNEAVLEWOLPVDVQNGFIRNYTIFRYIIGNETAIVNDSSHTEYTLSSL/SDTLY 897  
QY 20 -----DKHTCPPCPAPELLGGPSVFLFPPK 45  
DB 898 MVRMAAYTDEGGKDGPEFTTTPKAQGEIESGGDKTTCPCPAPELLGGPSVFLFPPK 957  
QY 46 PKDTLMISRTPEVTCVVVDVSHEDPEVKFNKYVDGVVHNATKPREQYNSTYRVSVL 105  
DB 958 PKDTLMISRTPEVTCVVVDVSHEDPEVKFNWYVDGVVHNATKPREQYNSTYRVSVL 1017

QY 106 TVLHODWLNKGEYKCKVSNKALPAPIEKTISKAKGQPREPQVYTLPPSRDELTKNQVSLT 165  
DB 1018 TVLHODWLNKGEYKCKVSNKALPAPIEKTISKAKGQPREPQVYTLPPSRDELTKNQVSLT 1077  
QY 166 CLVKGFPYSDIAVEWESNGQPENNYKTTTPVLDSDGSFFLYSKLTVDKSRWQQGNVFSCS 225  
DB 1078 CLVKGFPYSDIAVEWESNGQPENNYKTTTPVLDSDGSFFLYSKLTVDKSRWQQGNVFSCS 1137  
QY 226 VMHEALHNHYTQKSLSLSPGK 246  
DB 1138 VMHEALHNHYTQKSLSLSPGK 1158

Search completed: April 21, 2003, 10:45:22  
Job time : 32.9412 secs

GenCore version 5.1.4.p5.4578  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: April 21, 2003, 10:42:26 ; Search time 13.7471 seconds  
(without alignments)  
526.515 Million cell updates/sec

Title: 7LINK2

Perfect score: 1334

Sequence: 1 DWLKAIFYDKVAEKLEAFMD.....MHEALHNYTKSLSPGK 246

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.0

Searched: 262574 seqs, 29422922 residues

Total number of hits satisfying chosen parameters: 262574

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued\_Patents\_AA.\*

1: /cgn2\_6/ptodata/1/iaa/5A\_COMB.pep.\*

2: /cgn2\_6/ptodata/1/iaa/5B\_COMB.pep.\*

3: /cgn2\_6/ptodata/1/iaa/6A\_COMB.pep.\*

4: /cgn2\_6/ptodata/1/iaa/6B\_COMB.pep.\*

5: /cgn2\_6/ptodata/1/iaa/PCTUS\_COMB.pep.\*

6: /cgn2\_6/ptodata/1/iaa/backfiles1.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1258	94.3	476	3	US-08-487-550-12
2	1252	93.9	449	1	US-08-458-516-13
3	1250	93.7	476	2	US-08-378-939-10
4	1249	93.6	347	1	US-07-940-861-43
5	1249	93.6	347	1	US-08-459-512-43
6	1249	93.6	347	2	US-08-459-657-43
7	1249	93.6	347	2	US-08-460-132-43
8	1249	93.6	347	4	US-08-466-465-8
9	1249	93.6	347	5	PCT-US92-02050-43
10	1249	93.6	452	3	US-09-027-449-71
11	1249	93.6	452	4	US-09-026-985-71
12	1249	93.6	452	4	US-09-121-952A-71
13	1249	93.6	452	4	US-09-234-340A-71
14	1249	93.6	459	1	US-08-157-101A-7
15	1249	93.6	664	3	US-08-957-063-16
16	1249	93.6	664	3	US-08-957-063-18
17	1249	93.6	664	4	US-09-487-685-16
18	1249	93.6	664	4	US-09-487-685-18
19	1249	93.6	664	4	US-08-802-805D-16
20	1249	93.6	664	4	US-08-802-805D-18
21	1248	93.6	476	3	US-08-487-550-4
22	1247	93.5	437	5	PCT-US96-10043-11
23	1247	93.5	446	3	US-08-397-411-7
24	1246	93.4	454	2	US-07-934-373C-22
25	1246	93.4	454	3	US-08-437-642B-22
26	1246	93.4	454	4	US-08-146-206C-22
27	1246	93.4	454	5	PCT-US93-07832-22

28	1245	93.3	371	1	US-08-236-311-7	Sequence 7, Appl
29	1245	93.3	371	3	US-08-457-918-7	Sequence 7, Appl
30	1245	93.3	388	4	US-09-131-247-16	Sequence 16, Appl
31	1245	93.3	424	5	PCT-US95-03866-12	Sequence 12, Appl
32	1245	93.3	424	5	PCT-US95-03866-14	Sequence 14, Appl
33	1245	93.3	451	2	US-08-887-352B-14	Sequence 14, Appl
34	1245	93.3	451	2	US-08-887-352B-16	Sequence 16, Appl
35	1245	93.3	451	2	US-08-887-352B-18	Sequence 18, Appl
36	1245	93.3	451	3	US-08-466-151-65	Sequence 65, Appl
37	1245	93.3	451	4	US-09-109-207C-14	Sequence 14, Appl
38	1245	93.3	451	4	US-09-109-207C-16	Sequence 16, Appl
39	1245	93.3	451	4	US-09-109-207C-18	Sequence 18, Appl
40	1245	93.3	451	4	US-09-282-505-2	Sequence 2, Appl
41	1245	93.3	451	4	US-09-054-255-2	Sequence 2, Appl
42	1245	93.3	451	4	US-09-296-005-14	Sequence 14, Appl
43	1245	93.3	451	4	US-09-296-005-16	Sequence 16, Appl
44	1245	93.3	451	4	US-09-296-005-18	Sequence 18, Appl
45	1245	93.3	453	3	US-08-466-151-8	Sequence 8, Appl

ALIGNMENTS

RESULT 1  
US-08-487-550-12  
; Sequence 12, Application US/08487550  
; Patent No. 6113898  
; GENERAL INFORMATION:  
; APPLICANT: Anderson, Darrell R.  
; TITLE OF INVENTION: "MONKEY MONOCLONAL ANTIBODIES SPECIFIC  
; TITLE OF INVENTION: TO HUMAN B7.1 AND/OR B7.2 PRIMATIZED FORMS THEREOF,  
; TITLE OF INVENTION: PHARMACEUTICAL COMPOSITIONS CONTAINING, AND USE THEREOF AS  
; TITLE OF INVENTION: IMMUNOSUPPRESSANTS"  
; NUMBER OF SEQUENCES: 12  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: BURNS, DOANE, SWECKER & MATHIS  
; STREET: 699 Prince Street  
; CITY: Alexandria  
; STATE: VA  
; COUNTRY: USA  
; ZIP: 22314

COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.30  
CURRENT APPLICATION DATA: US/08/487,550  
APPLICATION NUMBER: US/08/487,550  
FILING DATE: 07-JUN-1995  
CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: Teskin, Robin L.  
REGISTRATION NUMBER: 35,030  
REFERENCE/DOCKET NUMBER: 012712-131  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 703-836-6620  
TELEFAX: 703-836-2021  
INFORMATION FOR SEQ ID NO: 12:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 476 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-487-550-12

Query Match 94.3%; Score 1258; DB 3; Length 476;  
Best Local Similarity 56.7%; Pred. No. 3.4e-44;  
Matches 233; Conservative 5; Mismatches 8; Indels 165; Gaps 1;  
QY 1 DWLKAIFYD-----  
:|: ||  
Db 66 EWIGSFYSSGNTYNPFLSKSVTISTDTSKNQFLKLSMTAADTAVYCVDRDLFSW 125

```
QY 9 ----- 8
Db 126 GMVNNWFDWGPGLVLTSSASTKGPSVFPLAPSKSTSGGTAALGCLVKDYFPEPTV 185
QY 9 -----KVAEKLK 15
Db 186 SWSNGALTSGVHTFPVAVLOSSGLYSLSVVTVPSSSLGTQTYICNVNHRPSNTKVDKXAE 245
QY 16 EAFMDKTHTCPPCPAPPELLGGPSVFLFPPKPKDTHLMSRTPEVTCVVVDVSHEDPEVKEN 75
Db 246 PKSCDKTHTCPPCPAPPELLGGPSVFLFPPKPKDTHLMSRTPEVTCVVVDVSHEDPEVKEN 305
QY 76 WYVDGVEVHNAKTPREEQYNSTYRVVSVLTVLHQDLNKGKEYCKKVSNAKALPAPIEKTI 135
Db 306 WYVDGVEVHNAKTPREEQYNSTYRVVSVLTVLHQDLNKGKEYCKKVSNAKALPAPIEKTI 365
QY 136 SKAGQPREPQVYTLPPSRDELTKNQVSLTCLVKGYGIPSDIAVEWESNGQPENNYKTTPP 195
Db 366 SKAGQPREPQVYTLPPSRDELTKNQVSLTCLVKGYGIPSDIAVEWESNGQPENNYKTTPP 425
QY 196 VLDSGDSFFLYSKLTVDKSRWQGNVFCSCVMHEALHNYHTOKSLSLSPGK 246
Db 426 VLDSGDSFFLYSKLTVDKSRWQGNVFCSCVMHEALHNYHTOKSLSLSPGK 476

RESULT 2
US-08-458-516-13
; Sequence 13, Application US/08458516
; Patent No. 5777085
; GENERAL INFORMATION:
; APPLICANT: Co, Man Sung
; APPLICANT: Tso, J. Yun
; TITLE OF INVENTION: Humanized Antibodies Reactive with
; TITLE OF INVENTION: GPIIB/IIIA
; NUMBER OF SEQUENCES: 23
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: William M. Smith
; STREET: One Market Plaza, Steuart Tower, Suite 2000
; CITY: San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94105
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/458,516
; FILING DATE:
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/059,159
; FILING DATE: 03-MAY-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: Smith, William M.
; REGISTRATION NUMBER: 30,223
; REFERENCE/DOCKET NUMBER: 11823-37-3
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415-326-2400
; TELEFAX: 415-326-2422
; INFORMATION FOR SEQ ID NO: 13:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 449 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-458-516-13

Query Match 93.9%; Score 1252; DB 1; Length 449;
Best Local Similarity 57.4%; Pred. No. 5.3e-44;
Matches 232; Conservative 5; Mismatches 9; Indels 158; Gaps 1;
```

```
QY 1 DMLKAFY ----- 7
Db 46 EMIGVLYPCSGGTNYNEKEFKGRVTLTVDESTNTAYMELSSLSRSEDPAVYFCARRDGNYGW 105
QY 8 ----- 7
Db 106 FAYWGGTGLTVTVSSASTKGPSVFPLAPSKSTSGGTAALGCLVKDYFPEPTVTVSMNSGAL 165
QY 8 -----DKVAEKLKAEAFMDKT 22
Db 166 TSGVHTFPVAVLOSSGLYSLSVVTVPSSSLGTQTYICNVNHRPSNTKVDKKEPKSCDKT 225
QY 23 HFCPPCPAPPELLGGPSVFLFPPKPKDTHLMSRTPEVTCVVVDVSHEDPEVKENWYVDGYE 82
Db 226 HFCPPCPAPPELLGGPSVFLFPPKPKDTHLMSRTPEVTCVVVDVSHEDPEVKENWYVDGYE 285
QY 83 VHNATKTPREEQYNSTYRVVSVLTVLHQDLNKGKEYCKKVSNAKALPAPIEKTIKAKGQP 142
Db 286 VHNATKTPREEQYNSTYRVVSVLTVLHQDLNKGKEYCKKVSNAKALPAPIEKTIKAKGQP 345
QY 143 REPQVYTLPPSRDELTKNQVSLTCLVKGYGIPSDIAVEWESNGQPENNYKTTPPVLDSDGS 202
Db 346 REPQVYTLPPSRDELTKNQVSLTCLVKGYGIPSDIAVEWESNGQPENNYKTTPPVLDSDGS 405
QY 203 FFYLSKLTVDKSRWQGNVFCSCVMHEALHNYHTOKSLSLSPGK 246
Db 406 FFYLSKLTVDKSRWQGNVFCSCVMHEALHNYHTOKSLSLSPGK 449

RESULT 3
US-08-378-939-10
; Sequence 10, Application US/08378939
; Patent No. 5876961
; GENERAL INFORMATION:
; APPLICANT: CROWE, JAMES SCOTT
; APPLICANT: LEWIS, ALAN PETER
; TITLE OF INVENTION: PRODUCTION OF ANTIBODIES
; NUMBER OF SEQUENCES: 46
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: ROTHWELL, FIGG, ERNST & KURZ
; STREET: 555 THIRTEENTH ST. N.W.
; CITY: WASHINGTON
; STATE: D. C.
; COUNTRY: U. S.
; ZIP: 20004
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/378,939
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/952640
; FILING DATE: 01-DEC-1992
; ATTORNEY/AGENT INFORMATION:
; NAME: ERNST, BARBARA G
; REGISTRATION NUMBER: 30,377
; REFERENCE/DOCKET NUMBER: 1808-118
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202) 783-6040
; TELEFAX: (202) 783-6031
; INFORMATION FOR SEQ ID NO: 10:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 476 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-378-939-10
```



```
Query Match 93.7%; Score 1250; DB 2; Length 476;
Best Local Similarity 56.6%; Pred. No. 7.4e-44;
Matches 233; Conservative 7; Mismatches 6; Indels 166; Gaps 2;

QY 1 DWL-----KAF 6
DB 65 EWMGIPLFGTPTYSQFQGRVITADKSTSTAHMELTSLRSEDPAVYVCATDRYQAN 124
QY 7 YD-----8
DB 125 FDRARVGFDPWCGTLVTVSSASTKGPSVFPPLAPSSKSTSGTAALGCLVKDYPPEVPT 184
QY 9 -----KVAEKL 14
DB 185 VSMNSGALTSGVHTFPAVLQSSGLYSLSSVTVTPSSSLGTQTQYICNVNHPKSNTKVDKVK 244
QY 15 KEAFMDKTHTCPCPAPELLGGPSVFLPPKPKDMLISRTPETVTCVVVDVSHEDPEVKF 74
DB 245 EPKSCDKTHTCPCPAPELLGGPSVFLPPKPKDMLISRTPETVTCVVVDVSHEDPEVKF 304
QY 75 NWYVDGVEVHNAKTKPREEQYNSTYRVVSVLTVLHODWLNKGEYCKVSNKALPAPIET 134
DB 305 NWYVDGVEVHNAKTKPREEQYNSTYRVVSVLTVLHODWLNKGEYCKVSNKALPAPIET 364
QY 135 ISKAKQPREPOVYTLPPSRDELTKNQVSLTCLVKGFYPSDIAVEWESNGQPENNYKTT 194
DB 365 ISKAKQPREPOVYTLPPSRDELTKNQVSLTCLVKGFYPSDIAVEWESNGQPENNYKTT 424
QY 195 PVLDSDGSFFLYSKLTVDKSRWQGNVFSCSVMEALHNNHYTKLSLSPGK 246
DB 425 PVLDSDGSFFLYSKLTVDKSRWQGNVFSCSVMEALHNNHYTKLSLSPGK 476

RESULT 4
US-07-940-861-43
; Sequence 43, Application US/07940861
; Patent No. 5547853
; GENERAL INFORMATION:
; APPLICANT: BIOGEN, INC.
; APPLICANT: WALLNER, Barbara P.
; APPLICANT: MILLER, Glenn T.
; APPLICANT: ROSA, Margaret D.
; TITLE OF INVENTION: CD2-BINDING DOMAIN OF LYMPHOCYTE
; NUMBER OF SEQUENCES: 43
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Fish & Neave
; STREET: 875 Third Avenue
; CITY: New York
; STATE: New York
; COUNTRY: U.S.A.
; ZIP: 10022-6250
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/940,861
; FILING DATE: 21-OCT-1992
; CLASSIFICATION: 514
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/US92/02050
; FILING DATE: 12-MAR-1992
; APPLICATION NUMBER: US 07/667,971
; FILING DATE: 12-MAR-1991
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/770,967
; FILING DATE: 07-OCT-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: HALEY, James F., Jr.
; REGISTRATION NUMBER: 27,794
```

```
REFERENCE/DOCKET NUMBER: B151CIP2
TELEPHONE: (212)715-0600
TELEFAX: (212)715-0673
TELEX: 14-8367
INFORMATION FOR SEQ ID NO: 43:
SEQUENCE CHARACTERISTICS:
LENGTH: 347 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-07-940-861-43

Query Match 93.6%; Score 1249; DB 1; Length 347;
Best Local Similarity 80.2%; Pred. No. 3.7e-44;
Matches 235; Conservative 1; Mismatches 8; Indels 49; Gaps 2;

QY 2 WLKAFDYKVAEKLKEAF-----18
DB 56 WKQK-KDKVAELENSEFRASFKNRVYLDTVSGSLTIYNLTSSDEDEYEMESPNITDTM 114
QY 19 -----MDKTHTCPCPAPELLGGPSVFLPPKPKDMLISRTPETVTCVVVDVSHEDPEVK 73
DB 115 KFFLYVDKTHTCPCPAPELLGGPSVFLPPKPKDMLISRTPETVTCVVVDVSHEDPEVK 174
QY 74 FMYVDGVEVHNAKTKPREEQYNSTYRVVSVLTVLHODWLNKGEYCKVSNKALPAPIEK 133
DB 175 FMYVDGVEVHNAKTKPREEQYNSTYRVVSVLTVLHODWLNKGEYCKVSNKALPAPIEK 234
QY 134 TISKAKQPREPOVYTLPPSRDELTKNQVSLTCLVKGFYPSDIAVEWESNGQPENNYKTT 193
DB 235 TISKAKQPREPOVYTLPPSRDELTKNQVSLTCLVKGFYPSDIAVEWESNGQPENNYKTT 294
QY 194 PVLDSDGSFFLYSKLTVDKSRWQGNVFSCSVMEALHNNHYTKLSLSPGK 246
DB 295 PVLDSDGSFFLYSKLTVDKSRWQGNVFSCSVMEALHNNHYTKLSLSPGK 347

RESULT 5
US-08-459-512-43
; Sequence 43, Application US/08459512
; Patent No. 5728677
; GENERAL INFORMATION:
; APPLICANT: BIOGEN, INC.
; APPLICANT: WALLNER, Barbara P.
; APPLICANT: MILLER, Glenn T.
; APPLICANT: ROSA, Margaret D.
; TITLE OF INVENTION: CD2-BINDING DOMAIN OF LYMPHOCYTE
; NUMBER OF SEQUENCES: 43
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Fish & Neave
; STREET: 875 Third Avenue
; CITY: New York
; STATE: New York
; COUNTRY: U.S.A.
; ZIP: 10022-6250
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/459,512
; FILING DATE: 02-JUN-1995
; CLASSIFICATION: 514
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/US92/02050
; FILING DATE: 12-MAR-1992
; APPLICATION NUMBER: US 07/667,971
; FILING DATE: 12-MAR-1991
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/770,967
```

;; FILING DATE: 07-OCT-1991  
;; ATTORNEY/AGENT INFORMATION:  
;; NAME: HALEY, James F., Jr.  
;; REGISTRATION NUMBER: 27,794  
;; REFERENCE/DOCKET NUMBER: B151CIP2  
;; TELECOMMUNICATION INFORMATION:  
;; TELEPHONE: (212)715-0600  
;; TELEFAX: (212)715-0673  
;; TELEX: 14-8367  
;; INFORMATION FOR SEQ ID NO: 43:  
;; SEQUENCE CHARACTERISTICS:  
;; LENGTH: 347 amino acids  
;; TYPE: amino acid  
;; TOPOLOGY: linear  
;; MOLECULE TYPE: protein  
US-08-459-512-43

Query Match 93.6%; Score 1249; DB 1; Length 347;

Best Local Similarity 80.2%; Pred. No. 3.7e-44;

Matches 235; Conservative 1; Mismatches 8; Indels 49; Gaps 2;

QY 2 WLKAFYDKVAEKLKEAF----- 18  
DB 56 WKQK-KDKVAELENSEFRASFKNRVYLDVTVSGSLTIYNTLSDEDEYEMESPNITDTM 114  
QY 19 -----MDKHTHTCPCPAPPELLGGPSVFLFPPKPKDTLMISRTPEVTCVVDVSHEDPEVK 73  
DB 115 KFFLYVDKTHTCPCPCAPPELLGGPSVFLFPPKPKDTLMISRTPEVTCVVDVSHEDPEVK 174  
QY 74 FMYVDGVEVHNKTKPREEQNSTYRVSVLTVLHQLDNLNGKEYCKVSKNKPAPTEK 133  
DB 175 FMYVDGVEVHNKTKPREEQNSTYRVSVLTVLHQLDNLNGKEYCKVSKNKPAPTEK 234  
QY 134 TISKAGGPPEQVYTLPPSRDELTKNOVSLTCLVKGFYPSDIAVEMESNGOPENNYKTT 193  
DB 235 TISKAGGPPEQVYTLPPSRDELTKNOVSLTCLVKGFYPSDIAVEMESNGOPENNYKTT 294  
QY 194 PPVLDSGSGFFLYSKLTVDKSRWQGNVFCSCVMHEALHNHYTKLSLSPCK 246  
DB 295 PPVLDSGSGFFLYSKLTVDKSRWQGNVFCSCVMHEALHNHYTKLSLSPCK 347

#### RESULT 6

US-08-459-657-43  
; Sequence 43, Application US/08459657  
; Patent No. 5914111

#### GENERAL INFORMATION:

;; APPLICANT: BIOGEN, INC.  
;; APPLICANT: WALLNER, Barbara P.  
;; APPLICANT: MILLER, Glenn T.  
;; APPLICANT: ROSA, Margaret D.  
;; TITLE OF INVENTION: CD2-BINDING DOMAIN OF LYMPHOCYTE  
;; TITLE OF INVENTION: FUNCTION ASSOCIATED ANTIGEN 3  
;; NUMBER OF SEQUENCES: 43  
;; CORRESPONDENCE ADDRESS:  
;; ADDRESSEE: Fish & Neave  
;; STREET: 875 Third Avenue  
;; CITY: New York  
;; STATE: New York  
;; COUNTRY: U.S.A.  
;; ZIP: 10022-6250

#### COMPUTER READABLE FORM:

;; MEDIUM TYPE: Floppy disk  
;; COMPUTER: IBM PC compatible  
;; OPERATING SYSTEM: PC-DOS/MS-DOS  
;; SOFTWARE: PatentIn Release #1.0, Version #1.25  
;; CURRENT APPLICATION DATA:  
;; APPLICATION NUMBER: US/08/459,657  
;; FILING DATE: 02-JUN-1995  
;; CLASSIFICATION: 424  
;; PRIOR APPLICATION DATA:  
;; APPLICATION NUMBER: PCT/US92/02050  
;; FILING DATE: 12-MAR-1992

;; APPLICATION NUMBER: US 07/667,971  
;; FILING DATE: 12-MAR-1991  
;; PRIOR APPLICATION DATA:  
;; APPLICATION NUMBER: US 07/770,967  
;; FILING DATE: 07-OCT-1991  
;; ATTORNEY/AGENT INFORMATION:  
;; NAME: HALEY, James F., Jr.  
;; REGISTRATION NUMBER: 27,794  
;; REFERENCE/DOCKET NUMBER: B151CIP2  
;; TELECOMMUNICATION INFORMATION:  
;; TELEPHONE: (212)715-0600  
;; TELEFAX: (212)715-0673  
;; TELEX: 14-8367  
;; INFORMATION FOR SEQ ID NO: 43:  
;; SEQUENCE CHARACTERISTICS:  
;; LENGTH: 347 amino acids  
;; TYPE: amino acid  
;; TOPOLOGY: linear  
;; MOLECULE TYPE: protein  
US-08-459-657-43

Query Match 93.6%; Score 1249; DB 2; Length 347;

Best Local Similarity 80.2%; Pred. No. 3.7e-44;

Matches 235; Conservative 1; Mismatches 8; Indels 49; Gaps 2;

QY 2 WLKAFYDKVAEKLKEAF----- 18  
DB 56 WKQK-KDKVAELENSEFRASFKNRVYLDVTVSGSLTIYNTLSDEDEYEMESPNITDTM 114  
QY 19 -----MDKHTHTCPCPAPPELLGGPSVFLFPPKPKDTLMISRTPEVTCVVDVSHEDPEVK 73  
DB 115 KFFLYVDKTHTCPCPCAPPELLGGPSVFLFPPKPKDTLMISRTPEVTCVVDVSHEDPEVK 174  
QY 74 FMYVDGVEVHNKTKPREEQNSTYRVSVLTVLHQLDNLNGKEYCKVSKNKPAPTEK 133  
DB 175 FMYVDGVEVHNKTKPREEQNSTYRVSVLTVLHQLDNLNGKEYCKVSKNKPAPTEK 234  
QY 134 TISKAGGPPEQVYTLPPSRDELTKNOVSLTCLVKGFYPSDIAVEMESNGOPENNYKTT 193  
DB 235 TISKAGGPPEQVYTLPPSRDELTKNOVSLTCLVKGFYPSDIAVEMESNGOPENNYKTT 294  
QY 194 PPVLDSGSGFFLYSKLTVDKSRWQGNVFCSCVMHEALHNHYTKLSLSPCK 246  
DB 295 PPVLDSGSGFFLYSKLTVDKSRWQGNVFCSCVMHEALHNHYTKLSLSPCK 347

#### RESULT 7

US-08-460-132-43  
; Sequence 43, Application US/08460132  
; Patent No. 5928643

#### GENERAL INFORMATION:

;; APPLICANT: BIOGEN, INC.  
;; APPLICANT: WALLNER, Barbara P.  
;; APPLICANT: MILLER, Glenn T.  
;; APPLICANT: ROSA, Margaret D.  
;; TITLE OF INVENTION: CD2-BINDING DOMAIN OF LYMPHOCYTE  
;; TITLE OF INVENTION: FUNCTION ASSOCIATED ANTIGEN 3  
;; NUMBER OF SEQUENCES: 43  
;; CORRESPONDENCE ADDRESS:  
;; ADDRESSEE: Fish & Neave  
;; STREET: 875 Third Avenue  
;; CITY: New York  
;; STATE: New York  
;; COUNTRY: U.S.A.  
;; ZIP: 10022-6250

#### COMPUTER READABLE FORM:

;; MEDIUM TYPE: Floppy disk  
;; COMPUTER: IBM PC compatible  
;; OPERATING SYSTEM: PC-DOS/MS-DOS  
;; SOFTWARE: PatentIn Release #1.0, Version #1.25  
;; CURRENT APPLICATION DATA:  
;; APPLICATION NUMBER: US/08/460,132  
;; FILING DATE: 02-JUN-1995

```

; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/940,861
; FILING DATE: 21-OCT-1992
; APPLICATION NUMBER: PCT/US92/02050
; FILING DATE: 12-MAR-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/667,971
; FILING DATE: 12-MAR-1991
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/770,967
; FILING DATE: 07-OCT-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: HALEY, James F., Jr.
; REGISTRATION NUMBER: 27,794
; REFERENCE/DOCKET NUMBER: B151CIP2
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212)715-0600
; TELEFAX: (212)715-0673
; TELEX: 14-8367
; INFORMATION FOR SEQ ID NO: 43:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 347 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-460-132-43

Query Match 93.6%; Score 1249; DB 2; Length 347;
Best Local Similarity 80.2%; Pred. No. 3.7e-44;
Matches 235; Conservative 1; Mismatches 8; Indels 49; Gaps 2;

QY 2 WLKAFYDKVAEKLKEAF-----18
| | | | |
Db 56 WKQK-KDKVAELENSEFRASFKNRYLDTVSGSLTYNLTSSDEDEYEMESPNTDTM 114

QY 19 -----MDKTHCTPCPAPELLGGPSVFLFPPKPKDTLMISRTPEVTCVVVDVSHEDPEVK 73
| | | | |
Db 115 KFFLYVDKTHCTPCPAPELLGGPSVFLFPPKPKDTLMISRTPEVTCVVVDVSHEDPEVK 174

QY 74 FNNYVDGVEVHNKTPREEQYNSTYRVVSVLTVLHQDWLNGKEYKCKVSNKALPAPIEK 133
| | | | |
Db 175 FNNYVDGVEVHNKTPREEQYNSTYRVVSVLTVLHQDWLNGKEYKCKVSNKALPAPIEK 234

QY 134 TISKAKQPREPQVYTLPPSRDELTKNOVSLTCLVKGFYPSDIAVEWESNGQPENNYKTT 193
| | | | |
Db 235 TISKAKQPREPQVYTLPPSRDELTKNOVSLTCLVKGFYPSDIAVEWESNGQPENNYKTT 294

QY 194 PPVLDSGSEFLYSLKLTVDKSRWQOGNVFSCVMHEALHNHYTQKSLSLSPGK 246
| | | | |
Db 295 PPVLDSGSEFLYSLKLTVDKSRWQOGNVFSCVMHEALHNHYTQKSLSLSPGK 347

RESULT 8
US-08-466-465-8
; Sequence 8, Application US/08466465
; Patent No. 6162432
; GENERAL INFORMATION:
; APPLICANT: Wallner, Barbara P.
; APPLICANT: Cooper, Kevin D.
; TITLE OF INVENTION: Method of Prophylaxis or Treatment of Antigen
; TITLE OF INVENTION: Presenting Cell Driven Skin Conditions Using
; TITLE OF INVENTION: Inhibitors of the CD2/LFA-3 Interaction
; NUMBER OF SEQUENCES: 8
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: LAHIVE & COCKFIELD
; STREET: 60 State Street, Suite 510
; City: Boston
; STATE: Massachusetts
; COUNTRY: USA
; ZIP: 02109-1875
; COMPUTER READABLE FORM: disk
; MEDIUM TYPE: Floppy disk

```

```

; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/466,465
; FILING DATE:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/US92/08755
; FILING DATE: 06-OCT-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/862,022
; FILING DATE: 12-APR-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/770,969
; FILING DATE: 07-OCT-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: Myers, Louis (PLM)
; REGISTRATION NUMBER: 35,965
; REFERENCE/DOCKET NUMBER: BGP-111CP
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617)227-7400
; TELEFAX: (617)227-5941
; INFORMATION FOR SEQ ID NO: 8:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 347 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-466-465-8

Query Match 93.6%; Score 1249; DB 4; Length 347;
Best Local Similarity 80.2%; Pred. No. 3.7e-44;
Matches 235; Conservative 1; Mismatches 8; Indels 49; Gaps 2;

QY 2 WLKAFYDKVAEKLKEAF-----18
| | | | |
Db 56 WKQK-KDKVAELENSEFRASFKNRYLDTVSGSLTYNLTSSDEDEYEMESPNTDTM 114

QY 19 -----MDKTHCTPCPAPELLGGPSVFLFPPKPKDTLMISRTPEVTCVVVDVSHEDPEVK 73
| | | | |
Db 115 KFFLYVDKTHCTPCPAPELLGGPSVFLFPPKPKDTLMISRTPEVTCVVVDVSHEDPEVK 174

QY 74 FNNYVDGVEVHNKTPREEQYNSTYRVVSVLTVLHQDWLNGKEYKCKVSNKALPAPIEK 133
| | | | |
Db 175 FNNYVDGVEVHNKTPREEQYNSTYRVVSVLTVLHQDWLNGKEYKCKVSNKALPAPIEK 234

QY 134 TISKAKQPREPQVYTLPPSRDELTKNOVSLTCLVKGFYPSDIAVEWESNGQPENNYKTT 193
| | | | |
Db 235 TISKAKQPREPQVYTLPPSRDELTKNOVSLTCLVKGFYPSDIAVEWESNGQPENNYKTT 294

QY 194 PPVLDSGSEFLYSLKLTVDKSRWQOGNVFSCVMHEALHNHYTQKSLSLSPGK 246
| | | | |
Db 295 PPVLDSGSEFLYSLKLTVDKSRWQOGNVFSCVMHEALHNHYTQKSLSLSPGK 347

RESULT 9
PCT-US92-02050-43
; Sequence 43, Application PC/TUS9202050
; GENERAL INFORMATION:
; APPLICANT: BIOGEN, INC.
; APPLICANT: WALLNER, Barbara P.
; APPLICANT: MILLER, Glenn T.
; APPLICANT: ROSA, Margaret D.
; TITLE OF INVENTION: CD2-BINDING DOMAIN OF LYMPHOCYTE
; TITLE OF INVENTION: FUNCTION ASSOCIATED ANTIGEN 3
; NUMBER OF SEQUENCES: 43
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Fish & Neave
; STREET: 875 Third Avenue
; City: New York
; STATE: New York
; COUNTRY: U.S.A.
; ZIP: 10022-6250

```

```

; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US92/02050
; FILING DATE: 19920312
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/667,971
; FILING DATE: 12-MAR-1991
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/770,967
; FILING DATE: 07-OCT-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: HALEY, James F., Jr.
; REGISTRATION NUMBER: 27,794
; REFERENCE/DOCKET NUMBER: B151CIP2
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212)715-0600
; TELEFAX: (212)715-0673
; TELEX: 14-8367
; INFORMATION FOR SEQ ID NO: 43:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 347 amino acids
; TYPE: AMINO ACID
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; PCT-US92-02050-43

Query Match 93.6%; Score 1249; DB 5; Length 347;
Best Local Similarity 80.2%; Pred. No. 3.7e-44;
Matches 235; Conservative 1; Mismatches 8; Indels 49; Gaps 2;

QY 2 WLKAFYDKVAEKLKEAF----- 18
DB 56 WKQK-KDKVAELENSEFRASFKNRYLDVTYVSGSLTIYNLTSSDEDEYEMESPNTIDPM 114
QY 19 -----MDKTHTCPPCPAPPELLGGPSVFLFPPPKPKDTLMISRTPEVTCVVDVSHEDPEVK 73
DB 115 KFELYVDKTHTCPPCPAPPELLGGPSVFLFPPPKPKDTLMISRTPEVTCVVDVSHEDPEVK 174
QY 74 FNYVDGVEVHNAKTPREEQNSTYRVSVTLVTHQDMLNGKEYKCKVSNKALPAPIEK 133
DB 175 FNYVDGVEVHNAKTPREEQNSTYRVSVTLVTHQDMLNGKEYKCKVSNKALPAPIEK 234
QY 134 TISKAGQPREPQVYTLPPSRDELTKNQVSLTCLVKGFYPSDIAVEWESNGQPENNYKT 193
DB 235 TISKAGQPREPQVYTLPPSRDELTKNQVSLTCLVKGFYPSDIAVEWESNGQPENNYKT 294
QY 194 PPVLDSGFSFLYSLKLTVDKSRWQGNVFCSCVMHEALHNHYTQKSLSLSPGK 246
DB 295 PPVLDSGFSFLYSLKLTVDKSRWQGNVFCSCVMHEALHNHYTQKSLSLSPGK 347

RESULT 10
US-09-027-449-71
; Sequence 71, Application US/09027449
; Patent No. 6025158
; GENERAL INFORMATION:
; APPLICANT: Gonzalez, Tania R.
; APPLICANT: Leong, Steven R.
; APPLICANT: Presta, Leonard G.
; TITLE OF INVENTION: Antibody Fragment-Polymer Conjugates and
; TITLE OF INVENTION: Humanized Anti-IL-8 Monoclonal Antibodies
; NUMBER OF SEQUENCES: 72
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Genentech, Inc.
; STREET: 1 DNA Way
; CITY: South San Francisco
; STATE: California
; COUNTRY: USA

```

```

; ZIP: 94080
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5 inch, 1.44 Mb floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: WinPatIn (Genentech)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/027,449
; FILING DATE: 20-Feb-1998
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 60/074,330
; FILING DATE: 22-Jan-1998
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 60/038,664
; FILING DATE: 21-Feb-1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Love, Richard B.
; REGISTRATION NUMBER: 34,659
; REFERENCE/DOCKET NUMBER: P1085R3-2
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 650/225-5530
; TELEFAX: 650/952-9881
; INFORMATION FOR SEQ ID NO: 71:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 452 amino acids
; TYPE: Amino Acid
; TOPOLOGY: Linear
; US-09-027-449-71

Query Match 93.6%; Score 1249; DB 3; Length 452;
Best Local Similarity 66.3%; Pred. No. 7.2e-44;
Matches 232; Conservative 6; Mismatches 5; Indels 107; Gaps 2;

QY 1 DWLKAFYD----- 8
DB 106 DW---FFDWGOGTLTVSSASTKGPSVFPAPSSKSTSGGTAALGCLVKDYFPEPTVTS 162
QY 9 -----KVAEKLKE 16
DB 163 WNSGALTSGVHTFPAVLQSSGLYSLSSVTVTPSSSLGTQTYICNVNHNKPSNTKVDKVEP 222
QY 17 AFMDKTHTCPPCPAPPELLGGPSVFLFPPPKPKDTLMISRTPEVTCVVDVSHEDPEVK 76
DB 223 KSCDKTHTCPPCPAPPELLGGPSVFLFPPPKPKDTLMISRTPEVTCVVDVSHEDPEVK 282
QY 77 YVDGVEVHNAKTPREEQNSTYRVSVTLVTHQDMLNGKEYKCKVSNKALPAPIEK 136
DB 283 YVDGVEVHNAKTPREEQNSTYRVSVTLVTHQDMLNGKEYKCKVSNKALPAPIEK 342
QY 137 KAKGQPREPQVYTLPPSRDELTKNQVSLTCLVKGFYPSDIAVEWESNGQPENNYKTTPV 196
DB 343 KAKGQPREPQVYTLPPSRDELTKNQVSLTCLVKGFYPSDIAVEWESNGQPENNYKTTPV 402
QY 197 LDSGSGFSFLYSLKLTVDKSRWQGNVFCSCVMHEALHNHYTQKSLSLSPGK 246
DB 403 LDSGSGFSFLYSLKLTVDKSRWQGNVFCSCVMHEALHNHYTQKSLSLSPGK 452

RESULT 11
US-09-026-985-71
; Sequence 71, Application US/09026985
; Patent No. 6133426
; GENERAL INFORMATION:
; APPLICANT: Gonzalez, Tania R.
; APPLICANT: Leong, Steven R.
; APPLICANT: Presta, Leonard G.
; TITLE OF INVENTION: Antibody Fragment-Polymer Conjugates and
; TITLE OF INVENTION: Humanized Anti-IL-8 Monoclonal Antibodies
; NUMBER OF SEQUENCES: 72
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Genentech, Inc.
; STREET: 1 DNA Way

```

```

; CITY: South San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94080
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5 inch, 1.44 Mb floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: WinPatIn (Genentech)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/026,985
; FILING DATE: 20-Feb-1998
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: Love, Richard B.
; REGISTRATION NUMBER: 34,659
; REFERENCE/DOCKET NUMBER: PI085R3-1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 650/225-5530
; TELEFAX: 650/952-9881
; INFORMATION FOR SEQ ID NO: 71:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 452 amino acids
; TYPE: Amino Acid
; TOPOLOGY: Linear
; US-09-026-985-71

Query Match 93.68; Score 1249; DB 4; Length 452;
Best Local Similarity 66.3%; Pred. No. 7.2e-44;
Matches 232; Conservative 6; Mismatches 5; Indels 107; Gaps 2;

Qy 1 DWLKAFYD----- 8
|||
Db 106 DW---FFDVWQGLTVTVSSASTKGPSVFPLAPSSKSTSGGTAALGCLVKDYFPEPVTVS 162
Qy 9 -----KVAEKLKE 16
|||
Db 163 WNSGALTSGVHTFPAVLQSSGLYSLSSVTVTPSSSLGTQTYICNVNHKPSNTKVDKKVEP 222
Qy 17 AFMDKTHTCPCPAPELLGGPSVFLFPPKPKDTLMISRTPEVTCVVDVSHEDPEVKFNW 76
Db 223 KSCDKTHTCPCPAPELLGGPSVFLFPPKPKDTLMISRTPEVTCVVDVSHEDPEVKFNW 282
Qy 77 YVDGVEVHNATKPREQYNSTYRVSVLTIVLHODWLNKGEYKCKVSNKALPAPIEKTIS 136
Db 283 YVDGVEVHNATKPREQYNSTYRVSVLTIVLHODWLNKGEYKCKVSNKALPAPIEKTIS 342
Qy 137 KAKGQPREPQVYTLPPPSRDELTKNOVSLTCLVKGFYPSDIAVEWESNGQPENNYKTTPPV 196
Db 343 KAKGQPREPQVYTLPPPSREEMTKNQVSLTCLVKGFYPSDIAVEWESNGQPENNYKTTPPV 402
Qy 197 LDSGSEFLYSLKLTVDKSRWQGNVFCSCVMHEALHNHYTOKSLSPGK 246
Db 403 LDSGSEFLYSLKLTVDKSRWQGNVFCSCVMHEALHNHYTOKSLSPGK 452

RESULT 12
US-09-121-952A-71
; Sequence 71, Application US/09121952A
; Patent No. 6458355
; GENERAL INFORMATION:
; APPLICANT: Genentech, Inc., Hsei, Vanessa
; APPLICANT: Koumenis, Iphigenia
; APPLICANT: Leong, Steven R.
; APPLICANT: Presta, Leonard G.
; APPLICANT: Shahrokhi, Zahra
; APPLICANT: Zapata, Gerardo A.
; TITLE OF INVENTION: METHODS OF TREATING INFLAMMATORY DISEASES
; WITH ANTI-IL-8 ANTIBODY FRAGMENT-POLYMER CONJUGATES
; NUMBER OF SEQUENCES: 72
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Genentech, Inc.
; STREET: 1 DNA Way
```

```

; CITY: South San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94080
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5 inch, 1.44 Mb floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: WinPatIn (Genentech)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/121,952A
; FILING DATE: 24-Jul-1998
; CLASSIFICATION: 514
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 60/074330
; FILING DATE: 22-JAN-1998
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 60/075467
; FILING DATE: 20-FEB-1998
; ATTORNEY/AGENT INFORMATION:
; NAME: Love, Richard B.
; REGISTRATION NUMBER: 34,659
; REFERENCE/DOCKET NUMBER: PI085R4
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 650/225-5530
; TELEFAX: 650/952-9881
; INFORMATION FOR SEQ ID NO: 71:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 452 amino acids
; TYPE: Amino Acid
; TOPOLOGY: Linear
; US-09-121-952A-71

Query Match 93.68; Score 1249; DB 4; Length 452;
Best Local Similarity 66.3%; Pred. No. 7.2e-44;
Matches 232; Conservative 6; Mismatches 5; Indels 107; Gaps 2;

Qy 1 DWLKAFYD----- 8
|||
Db 106 DW---FFDVWQGLTVTVSSASTKGPSVFPLAPSSKSTSGGTAALGCLVKDYFPEPVTVS 162
Qy 9 -----KVAEKLKE 16
|||
Db 163 WNSGALTSGVHTFPAVLQSSGLYSLSSVTVTPSSSLGTQTYICNVNHKPSNTKVDKKVEP 222
Qy 17 AFMDKTHTCPCPAPELLGGPSVFLFPPKPKDTLMISRTPEVTCVVDVSHEDPEVKFNW 76
Db 223 KSCDKTHTCPCPAPELLGGPSVFLFPPKPKDTLMISRTPEVTCVVDVSHEDPEVKFNW 282
Qy 77 YVDGVEVHNATKPREQYNSTYRVSVLTIVLHODWLNKGEYKCKVSNKALPAPIEKTIS 136
Db 283 YVDGVEVHNATKPREQYNSTYRVSVLTIVLHODWLNKGEYKCKVSNKALPAPIEKTIS 342
Qy 137 KAKGQPREPQVYTLPPPSRDELTKNOVSLTCLVKGFYPSDIAVEWESNGQPENNYKTTPPV 196
Db 343 KAKGQPREPQVYTLPPPSREEMTKNQVSLTCLVKGFYPSDIAVEWESNGQPENNYKTTPPV 402
Qy 197 LDSGSEFLYSLKLTVDKSRWQGNVFCSCVMHEALHNHYTOKSLSPGK 246
Db 403 LDSGSEFLYSLKLTVDKSRWQGNVFCSCVMHEALHNHYTOKSLSPGK 452

RESULT 13
US-09-234-340A-71
; Sequence 71, Application US/09234340A
; Patent No. 6468532
; GENERAL INFORMATION:
; APPLICANT: Genentech, Inc., Hsei, Vanessa
; APPLICANT: Koumenis, Iphigenia
; APPLICANT: Leong, Steven R.
; APPLICANT: Presta, Leonard G.
; APPLICANT: Shahrokhi, Zahra
; APPLICANT: Zapata, Gerardo A.
```

;; TITLE OF INVENTION: METHODS OF TREATING INFLAMMATORY DISEASES  
;; TITLE OF INVENTION: WITH ANTI-IL-8 ANTIBODY FRAGMENT-POLYMER CONJUGATES  
;; NUMBER OF SEQUENCES: 72  
;; CORRESPONDENCE ADDRESS:  
;; ADDRESSEE: Genentech, Inc.  
;; STREET: 1 DNA Way  
;; CITY: South San Francisco  
;; STATE: California  
;; COUNTRY: USA  
;; ZIP: 94080  
;; COMPUTER READABLE FORM:  
;; MEDIUM TYPE: 3.5 inch, 1.44 Mb floppy disk  
;; COMPUTER: IBM PC compatible  
;; OPERATING SYSTEM: PC-DOS/MS-DOS  
;; SOFTWARE: WinPatIn (Genentech)  
;; CURRENT APPLICATION DATA:  
;; APPLICATION NUMBER: US/09/234,340A  
;; FILING DATE:  
;; CLASSIFICATION:  
;; PRIOR APPLICATION DATA:  
;; APPLICATION NUMBER: US/09/121,952  
;; FILING DATE: 24-Jul-1998  
;; APPLICATION NUMBER: 60/074330  
;; FILING DATE: 22-JAN-1998  
;; PRIOR APPLICATION DATA:  
;; APPLICATION NUMBER: 60/075467  
;; FILING DATE: 20-FEB-1998  
;; ATTORNEY/AGENT INFORMATION:  
;; NAME: Love, Richard B.  
;; REGISTRATION NUMBER: 34,659  
;; REFERENCE/DOCKET NUMBER: P1085R4  
;; TELECOMMUNICATION INFORMATION:  
;; TELEPHONE: 650/225-5530  
;; TELEFAX: 650/952-9881  
;; INFORMATION FOR SEQ ID NO: 71:  
;; SEQUENCE CHARACTERISTICS:  
;; LENGTH: 452 amino acids  
;; TYPE: Amino Acid  
;; TOPOLOGY: Linear  
US-09-234-340A-71  
  
Query Match 93.6%; Score 1249; DB 4; Length 452;  
Best Local Similarity 66.3%; Pred. No. 7.2e-44;  
Matches 232; Conservative 6; Mismatches 5; Indels 107; Gaps 2;  
  
QY 1 DWLKAFYD-----KVAEKLKE 16  
|| ||  
Db 106 DW---FFDVGWGGLTVTVSSASTKGPSVFPPLAPSSKSTSGGTAALGCLVKDYFPEPTVYS 162  
QY 9 -----KVAEKLKE 16  
|| ||  
Db 163 WNSGALTSVHTFPAVLQSSGLYSLSSVTVVPSISLGTQTYICNVNHRKPSNTKVDKVEP 222  
QY 17 AFMDKTHTCPPCAPPELLGGPSVFLFPPPKDPTLMISRTPEVTCVVDVSHEDPEVKRW 76  
Db 223 KSCDKHTHTCPPCAPPELLGGPSVFLFPPPKDPTLMISRTPEVTCVVDVSHEDPEVKRW 282  
QY 77 YVDGVEVHNATKPREQYNSTYRVVSVLTVLHQDWLNGKEYCKVSNKALPAPTEKTI 136  
Db 283 YVDGVEVHNATKPREQYNSTYRVVSVLTVLHQDWLNGKEYCKVSNKALPAPTEKTI 342  
QY 137 KAGQPREPQVYTLPPSRDELTKNQVSLTCLVKGFYPSDIAVEWESNGQPENNYKTTTPV 196  
Db 343 KAGQPREPQVYTLPPSRDEMTKNQVSLTCLVKGFYPSDIAVEWESNGQPENNYKTTTPV 402  
QY 197 LDSGGSFFLYSKLTVDKSRWQQGNVFSCSVMEALHNHYTQKSLSLSPGK 246  
Db 403 LDSGGSFFLYSKLTVDKSRWQQGNVFSCSVMEALHNHYTQKSLSLSPGK 452  
  
RESULT 14  
US-08-157-101A-7  
; .Sequence 7, Application US/08157101A

;; Patent No. 5808032  
;; GENERAL INFORMATION:  
;; APPLICANT: KURIHARA, TATSUYA  
;; APPLICANT: MATSUKURA, SHIGEKAZU  
;; APPLICANT: TSURUOKA, NOBUO  
;; APPLICANT: ARIMA, KENJI  
;; APPLICANT: NISHIHARA, TATSURO  
;; TITLE OF INVENTION: ANTI-HBS ANTIBODY GENES AND EXPRESSION  
;; TITLE OF INVENTION: PLASMIDS THEREFOR  
;; NUMBER OF SEQUENCES: 9  
;; CORRESPONDENCE ADDRESS:  
;; ADDRESSEE: PILLSBURY, MADISON & SUTRO  
;; STREET: 1100 NEW YORK AVENUE, N.W.  
;; CITY: WASHINGTON  
;; STATE: D.C.  
;; COUNTRY: USA  
;; ZIP: 20005  
;; COMPUTER READABLE FORM:  
;; MEDIUM TYPE: Floppy disk  
;; COMPUTER: IBM PC compatible  
;; OPERATING SYSTEM: PC-DOS/MS-DOS  
;; SOFTWARE: PatentIn Release #1.0, Version #1.25  
;; CURRENT APPLICATION DATA:  
;; APPLICATION NUMBER: US/08/157,101A  
;; FILING DATE: 05-APR-1994  
;; CLASSIFICATION: 530  
;; ATTORNEY/AGENT INFORMATION:  
;; NAME: TITUS, MARLANA K  
;; REGISTRATION NUMBER: 35843  
;; REFERENCE/DOCKET NUMBER: 9437/204199  
;; TELECOMMUNICATION INFORMATION:  
;; TELEPHONE: 202-861-3711  
;; TELEFAX: 202-822-0944  
;; TELEX: 6714627 CUCH  
;; INFORMATION FOR SEQ ID NO: 7:  
;; SEQUENCE CHARACTERISTICS:  
;; LENGTH: 459 amino acids  
;; TYPE: amino acid  
;; STRANDEDNESS: single  
;; TOPOLOGY: linear  
;; MOLECULE TYPE: peptide  
US-08-157-101A-7  
  
Query Match 93.6%; Score 1249; DB 1; Length 459;  
Best Local Similarity 96.6%; Pred. No. 7.4e-44;  
Matches 230; Conservative 3; Mismatches 5; Indels 0; Gaps 0;  
  
QY 9 KVAEKLKEAFMDKTHTCPPCAPPELLGGPSVFLFPPPKDPTLMISRTPEVTCVVDVSH 68  
|| ||  
Db 222 KVDKKEVPKSCDKHTHTCPPCAPPELLGGPSVFLFPPPKDPTLMISRTPEVTCVVDVSH 281  
QY 69 DPEVKFNHYVDGVEVHNATKPREQYNSTYRVVSVLTVLHQDWLNGKEYCKVSNKALP 128  
Db 282 DPEVKFNHYVDGVEVHNATKPREQYNSTYRVVSVLTVLHQDWLNGKEYCKVSNKALP 341  
QY 129 APIETISKAGQPREPQVYTLPPSRDELTKNQVSLTCLVKGFYPSDIAVEWESNGQ 188  
Db 342 APIETISKAGQPREPQVYTLPPSRDELTKNQVSLTCLVKGFYPSDIAVEWESNGQ 401  
QY 189 NYKTPPPVLDSDGSFFLYSKLTVDKSRWQQGNVFSCSVMEALHNHYTQKSLSLSPGK 246  
Db 402 NYKTPPPVLDSDGSFFLYSKLTVDKSRWQQGNVFSCSVMEALHNHYTQKSLSLSPGK 459  
  
RESULT 15  
US-08-957-063-16  
; .Sequence 16, Application US/08957063  
;; Patent No. 6025157  
;; GENERAL INFORMATION:  
;; APPLICANT: Robert D. Klein, Arnon Rosenthal, Mary A. Hynes  
;; TITLE OF INVENTION: Neurturin Receptor  
;; NUMBER OF SEQUENCES: 19  
;; CORRESPONDENCE ADDRESS:

ADDRESSEE: Genentech, Inc.  
STREET: 1 DNA Way  
CITY: South San Francisco  
STATE: California  
COUNTRY: USA  
ZIP: 94080  
COMPUTER READABLE FORM:  
MEDIUM TYPE: 3.5 inch, 1.44 Mb floppy disk  
OPERATING SYSTEM: IBM PC compatible  
SOFTWARE: WinPatIn (Genentech)  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: 913  
FILING DATE: 18-Feb-1997  
FILING DATE: 24-Oct-1997  
CLASSIFICATION: 800  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/871  
FILING DATE: 9-Jun-1997  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 913  
FILING DATE: 18-Feb-1997  
ATTORNEY/AGENT INFORMATION:  
NAME: Torchia, PhD., Timothy E.  
REGISTRATION NUMBER: 36,700  
REFERENCE/DOCKET NUMBER: P1086P2  
TELEPHONE: 650/225-8674  
TELEFAX: 650/952-9881  
INFORMATION FOR SEQ ID NO: 16:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 664 amino acids  
TYPE: Amino Acid  
TOPOLOGY: Linear  
US-08-957-063-16

Query Match 93.6%; Score 1249; DB 3; Length 664;  
Best Local Similarity 41.8%; Pred. No. 1.9e-43;  
Matches 232; Conservative 9; Mismatches 4; Indels 310; Gaps 3;  
QY 2 W-----2  
DB 110 WSHLGLTEGEFEYASPYEPTVTSRLSDIFRLASIFSGTGADPVVSAKSNHCLDAKACN 169  
QY 3 -----LKAFYDKV-----10  
DB 170 LNDNCKKLRSYISICNREISPTERCNCRKCHALROFFDRVPSEYTYRMLFCSCDOAC 229  
QY 11 AEKLKEAFM-----19  
DB 230 AERRQTILPSCSYEDKEKPNCLDLRGVCRTHLCRLADFHANCRASTVTSCPADN 289  
QY 20 -----19  
DB 290 YQACLSYAGMIGDMPNVDSSPTGIVVPMCSGSGNMEEECEKFLRDTENPCLR 349  
QY 20 -----19  
DB 350 NAIQAFGNGTDVNVSPKGPSFQATQAPRVEKTPSLPDDLSDSTSLGTSVITCTSVQEQ 409  
QY 20 -----DKTHCTCPGPAPELLGGPSVFLFPPKPKDTLM 51  
DB 410 LKANNKSLSMCFELTTNIIPGRDPVDKTHCTCPGPAPELLGGPSVFLFPPKPKDTLM 469  
QY 52 ISRTPEVTCVVVDVSHEDPEVKFNWYVDGVEVHNAKTKPREEQYNSTYRVSVLTVHLQD 111  
DB 470 ISRTPEVTCVVVDVSHEDPEVKFNWYVDGVEVHNAKTKPREEQYNSTYRVSVLTVHLQD 529  
QY 112 WLNKKEYCKVSKNALKAPAEKTSKAKGPQRPQVYTLPPSRDELTKNOVSLTCLVKGF 171  
DB 530 WLNKKEYCKVSKNALKAPAEKTSKAKGPQRPQVYTLPPSRDELTKNOVSLTCLVKGF 589  
QY 172 YPSDIAVWESNGQPENNYKTTPPVLDSDGSFFLYSKLTVDKSRWQQGNVFSCSVMHEAL 231

Db 590 YPSDIAVWESNGQPENNYKTTPPVLDSDGSFFLYSKLTVDKSRWQQGNVFSCSVMHEAL 649  
QY 232 HHHTQKSLSLSPGK 246  
Db 650 HHHTQKSLSLSPGK 664  
Search completed: April 21, 2003, 10:50:23  
Job time : 15.7471 secs





GenCore version 5.1.4.p5.4578  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: April 21, 2003, 10:43:21 ; Search time 11.3353 Seconds  
(without alignments)

1640.982 Million cell updates/sec

Title: 7LINK2

Perfect score: 1334

Sequence: 1 DWLKFYDKVAEKLEAFMD.....MHEALHHVTKLSLSLSPGK 246

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.0

Searched: 288829 seqs, 7561385 residues

Total number of hits satisfying chosen parameters: 288829

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Published Applications AA:\*

- 1: /cgn2\_6/ptodata/2/pubpaa/US08\_NEW\_PUB.pep.\*
- 2: /cgn2\_6/ptodata/2/pubpaa/PCT\_NEW\_PUB.pep.\*
- 3: /cgn2\_6/ptodata/2/pubpaa/US06\_NEW\_PUB.pep.\*
- 4: /cgn2\_6/ptodata/2/pubpaa/US06\_PUBCOMB.pep.\*
- 5: /cgn2\_6/ptodata/2/pubpaa/US07\_NEW\_PUB.pep.\*
- 6: /cgn2\_6/ptodata/2/pubpaa/US07\_PUBCOMB.pep.\*
- 7: /cgn2\_6/ptodata/2/pubpaa/PCTUS\_PUBCOMB.pep.\*
- 8: /cgn2\_6/ptodata/2/pubpaa/US08\_PUBCOMB.pep.\*
- 9: /cgn2\_6/ptodata/2/pubpaa/US09\_NEW\_PUB.pep.\*
- 10: /cgn2\_6/ptodata/2/pubpaa/US09\_PUBCOMB.pep.\*
- 11: /cgn2\_6/ptodata/2/pubpaa/US10\_NEW\_PUB.pep.\*
- 12: /cgn2\_6/ptodata/2/pubpaa/US10\_PUBCOMB.pep.\*
- 13: /cgn2\_6/ptodata/2/pubpaa/US60\_NEW\_PUB.pep.\*
- 14: /cgn2\_6/ptodata/2/pubpaa/US60\_PUBCOMB.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	1258	94.3	476	9	US-10-124-905-12
2	1258	94.3	476	9	US-09-948-429B-12
3	1257	94.2	951	9	US-09-935-868-7
4	1254	94.0	859	9	US-09-935-868-7
5	1254	94.0	1158	9	US-09-935-868-26
6	1254	94.0	1168	9	US-09-935-868-24
7	1253	93.9	567	10	US-09-825-561A-16
8	1251	93.8	475	10	US-09-740-002-27
9	1251	93.8	476	10	US-09-747-669-3
10	1250	93.7	399	9	US-09-832-659-2
11	1250	93.7	423	9	US-09-832-659-44
12	1249	93.6	347	9	US-10-091-236-17
13	1249	93.6	347	9	US-10-091-313-7
14	1249	93.6	347	9	US-10-091-268-7
15	1249	93.6	347	10	US-09-796-033-8
16	1249	93.6	347	10	US-09-730-465-8
17	1249	93.6	452	9	US-09-726-258-71
18	1249	93.6	475	10	US-09-740-002-25
19	1249	93.6	547	10	US-09-746-359A-54

20	1249	93.6	571	10	US-09-746-359A-53
21	1249	93.6	594	10	US-09-815-108-22
22	1249	93.6	731	10	US-09-825-012-46
23	1249	93.6	741	10	US-09-825-012-55
24	1248	93.6	418	9	US-09-832-659-42
25	1248	93.6	476	9	US-10-124-905-4
26	1248	93.6	476	9	US-09-948-429B-4
27	1248	93.6	776	9	US-09-935-868-36
28	1248	93.6	776	9	US-09-935-868-40
29	1248	93.6	776	9	US-09-935-868-44
30	1248	93.6	910	9	US-09-935-868-28
31	1246	93.4	380	10	US-09-948-018-39
32	1246	93.4	451	10	US-09-822-698A-26
33	1245	93.3	388	10	US-09-784-623-16
34	1245	93.3	451	9	US-09-925-179-65
35	1245	93.3	451	9	US-09-925-179-66
36	1245	93.3	451	9	US-09-925-179-68
37	1245	93.3	451	10	US-09-920-171-14
38	1245	93.3	451	10	US-09-920-171-16
39	1245	93.3	451	10	US-09-920-171-18
40	1245	93.3	453	9	US-09-925-179-8
41	1245	93.3	453	10	US-09-802-077-8
42	1245	93.3	453	10	US-09-802-096-8
43	1245	93.3	691	9	US-09-935-868-20
44	1245	93.3	694	9	US-09-935-868-22
45	1245	93.3	780	9	US-09-935-868-34

#### ALIGNMENTS

#### RESULT 1

US-10-124-905-12

; Sequence 12, Application US/10124905

; Patent No. US20020166136A1

; GENERAL INFORMATION:

; APPLICANT: Anderson, Darrell R.

; TITLE OF INVENTION: "MONKEY MONOCLONAL ANTIBODIES SPECIFIC

; TO HUMAN B7.1 AND/OR B7.2 PRIMATIZED FORMS THEREOF,

; TITLE OF INVENTION: PHARMACEUTICAL COMPOSITIONS CONTAINING, AND USE THEREOF AS

; IMMUNOSUPPRESSANTS"

; NUMBER OF SEQUENCES: 12

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: BURNS, DOANE, SWECKER & MATHIS

; STREET: 699 Prince Street

; CITY: Alexandria

; STATE: VA

; COUNTRY: USA

; ZIP: 22314

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: PatentIn Release #1.0, Version #1.30

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/10/124,905

; FILING DATE:

; CLASSIFICATION:

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: 09/383,916

; FILING DATE:

; APPLICATION NUMBER: US 08/487,550

; FILING DATE: 07-JUN-1995

; ATTORNEY/AGENT INFORMATION:

; NAME: Teskin, Robin L.

; REGISTRATION NUMBER: 35,030

; REFERENCE/DOCKET NUMBER: 012712-131

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: 703-836-6620

; TELEFAX: 703-836-2021

; INFORMATION FOR SEQ ID NO: 12:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 476 amino acids

Sequence 53, Appl  
Sequence 22, Appl  
Sequence 46, Appl  
Sequence 55, Appl  
Sequence 42, Appl  
Sequence 4, Appl  
Sequence 36, Appl  
Sequence 40, Appl  
Sequence 28, Appl  
Sequence 39, Appl  
Sequence 26, Appl  
Sequence 16, Appl  
Sequence 16, Appl  
Sequence 65, Appl  
Sequence 66, Appl  
Sequence 68, Appl  
Sequence 14, Appl  
Sequence 16, Appl  
Sequence 18, Appl  
Sequence 8, Appl  
Sequence 8, Appl  
Sequence 8, Appl  
Sequence 20, Appl  
Sequence 22, Appl  
Sequence 34, Appl

```

; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-10-124-905-12

Query Match          94.3%; Score 1258; DB 9; Length 476;
Best Local Similarity 56.7%; Pred. No. 7.1e-33;
Matches 233; Conservative 5; Mismatches 8; Indels 165; Gaps 1;

QY 1 DWLKAFYD----- 8
   ||: ||
Db 66 EWIGSFYSSSGNTYYNPSLKQSVTISTDTSKNOFSLKLSMTAADTAVYYCVDRRLFV 125
QY 9 ----- 8
Db 126 GWYNNWFDVWGPGLVTVSSASTKGPSVFPPLAPSSKSTSGGTAALGCLVKDYFPEPTV 185
QY 9 -----KVAEKLK 15
   ||: ||
Db 186 SWNSGALTSGVHTFPVAVLQSSGLYSLSSVTVTPSSSLGTQTYICNVNHNKPSNTKVDKAE 245
QY 16 EAFMDKTHTCPCPCAPPELLGGPSVFLFPPPKDPTLMISRTPEVTCVVVDVSHEDPEVKEN 75
Db 246 PKSCDKTHTCPCPCAPPELLGGPSVFLFPPPKDPTLMISRTPEVTCVVVDVSHEDPEVKEN 305
QY 76 WYVDGVEVHNAKTKPREEQYNSTYRVVSVLTVLHQDLNKGKEYKCKVSNKALPAPIEKT 135
Db 306 WYVDGVEVHNAKTKPREEQYNSTYRVVSVLTVLHQDLNKGKEYKCKVSNKALPAPIEKT 365
QY 136 SKAKGQPREPQVYTLPPSRDELTKNQVSLTCLVKGYGSPIDIAVEWESNGQPENNYKTTTP 195
Db 366 SKAKGQPREPQVYTLPPSRDELTKNQVSLTCLVKGYGSPIDIAVEWESNGQPENNYKTTTP 425
QY 196 VLDSGSEFFLYSLKLTVDKSRWQGNVFCVSMHEALHNHYTQKSLSLSPGK 246
Db 426 VLDSGSEFFLYSLKLTVDKSRWQGNVFCVSMHEALHNHYTQKSLSLSPGK 476

RESULT 2
US-09-948-429B-12
; Sequence 12, Application US/09948429B
; Patent No. US20020177689A1
; GENERAL INFORMATION:
; APPLICANT: Anderson, Darrell R.
; TITLE OF INVENTION: "MONKEY MONOCLONAL ANTIBODIES SPECIFIC
; TO HUMAN B7.1 AND/OR B7.2 PRIMATIZED FORMS THEREOF."
; TITLE OF INVENTION: PHARMACEUTICAL COMPOSITIONS CONTAINING, AND USE THEREOF AS
; IMMUNOSUPPRESSANTS"
; NUMBER OF SEQUENCES: 12
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: BURNS, DOANE, SWECKER & MATHIS
; STREET: 699 Prince Street
; CITY: Alexandria
; STATE: VA
; COUNTRY: USA
; ZIP: 22314
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/948,429B
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 09/383,916
; FILING DATE:
; APPLICATION NUMBER: US 08/487,550
; FILING DATE: 07-JUN-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Teskin, Robin L.
; REGISTRATION NUMBER: 35,030
```

```

; REFERENCE/DOCKET NUMBER: 012712-131
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 703-836-6620
; TELEFAX: 703-836-2021
; INFORMATION FOR SEQ ID NO: 12:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 476 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-09-948-429B-12

Query Match          94.3%; Score 1258; DB 9; Length 476;
Best Local Similarity 56.7%; Pred. No. 7.1e-33;
Matches 233; Conservative 5; Mismatches 8; Indels 165; Gaps 1;

QY 1 DWLKAFYD----- 8
   ||: ||
Db 66 EWIGSFYSSSGNTYYNPSLKQSVTISTDTSKNOFSLKLSMTAADTAVYYCVDRRLFV 125
QY 9 ----- 8
Db 126 GWYNNWFDVWGPGLVTVSSASTKGPSVFPPLAPSSKSTSGGTAALGCLVKDYFPEPTV 185
QY 9 -----KVAEKLK 15
   ||: ||
Db 186 SWNSGALTSGVHTFPVAVLQSSGLYSLSSVTVTPSSSLGTQTYICNVNHNKPSNTKVDKAE 245
QY 16 EAFMDKTHTCPCPCAPPELLGGPSVFLFPPPKDPTLMISRTPEVTCVVVDVSHEDPEVKEN 75
Db 246 PKSCDKTHTCPCPCAPPELLGGPSVFLFPPPKDPTLMISRTPEVTCVVVDVSHEDPEVKEN 305
QY 76 WYVDGVEVHNAKTKPREEQYNSTYRVVSVLTVLHQDLNKGKEYKCKVSNKALPAPIEKT 135
Db 306 WYVDGVEVHNAKTKPREEQYNSTYRVVSVLTVLHQDLNKGKEYKCKVSNKALPAPIEKT 365
QY 136 SKAKGQPREPQVYTLPPSRDELTKNQVSLTCLVKGYGSPIDIAVEWESNGQPENNYKTTTP 195
Db 366 SKAKGQPREPQVYTLPPSRDELTKNQVSLTCLVKGYGSPIDIAVEWESNGQPENNYKTTTP 425
QY 196 VLDSGSEFFLYSLKLTVDKSRWQGNVFCVSMHEALHNHYTQKSLSLSPGK 246
Db 426 VLDSGSEFFLYSLKLTVDKSRWQGNVFCVSMHEALHNHYTQKSLSLSPGK 476

RESULT 3
US-09-935-868-9
; Sequence 9, Application US/09935868
; Patent No. US20020164690A1
; GENERAL INFORMATION:
; APPLICANT: Regeneron Pharmaceuticals, Inc
; TITLE OF INVENTION: Receptor Based Antagonists, and Methods of Making and Using
; FILE REFERENCE: REG 203D
; CURRENT APPLICATION NUMBER: US/09/935,868
; CURRENT FILING DATE: 2002-04-11
; PRIOR APPLICATION NUMBER: PCT/US99/22045
; PRIOR FILING DATE: 1999-09-22
; NUMBER OF SEQ ID NOS: 52
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 9
; LENGTH: 951
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-935-868-9

Query Match          94.2%; Score 1257; DB 9; Length 951;
Best Local Similarity 37.0%; Pred. No. 3.7e-32;
Matches 237; Conservative 4; Mismatches 5; Indels 394; Gaps 3;

QY 1 DW----- 2
   ||
Db 312 DMSEASGITYEDRPSKAPSFWKIDPSHTQGYRTVQLVWKTLPPFEANGKILDYEVTLT 371
```

```
Qy 3 -----LK 4
Db 372 RWKSHLQNTVNAKLTNLNDRYLATLTVRNLVGKSDAAVLITPACDFQATHPMDLK 431
Qy 5 AF-----6
Db 432 APPKDNMLWENTTPRESVKYILEWCVLSDKAPCIITDQOEDGTVHRTYLRGNLAESKC 491
Qy 7 -----YD-----8
Db 492 YLITVPVYADPGSPESIKAYLKQAPSGKGTVRTKKVGNKNEAVLEWDQLPVDVQNGFI 551
Qy 9 -----8
Db 552 RNYTIFYRTIICNETAVNVDSSTHTYTLTSLTSDTLVMRMAAYTDEGGKDGPEFTFTTP 611
Qy 9 -----8
Db 612 KPAQGEIESGASTKGPSVFPLAPLSKSTSGGTAALGCLVKDYFPEPTVTSNNSGALTSGV 671
Qy 9 -----KVAEKLKEAFMDKTHTCP 26
Db 672 HTFPAVLQSSGLYSLSVVTPSSSLGTQTYICNVNHKPSNTKVDKKVEPKSCDKTHTCP 731
Qy 27 PCPAPELLGGPSVFLFPPKPKDTLMISRTPEVTCVVDVSHEDPEVKFNWYVDGVVEVHNA 86
Db 732 PCPAPELLGGPSVFLFPPKPKDTLMISRTPEVTCVVDVSHEDPEVKFNWYVDGVVEVHNA 791
Qy 87 KTKPREEQYNSTYRVVSVLTVLHODWLNKGEYCKVSKNKPALPAPIEKTISKAKGQPREPQ 146
Db 792 KTKPREEQYNSTYRVVSVLTVLHODWLNKGEYCKVSKNKPALPAPIEKTISKAKGQPREPQ 851
Qy 147 VYTLPPSRDELTKNQVSLTCLVKGYPSDIAVEWESNGQPENNYKTTTPVLDSDGSFFLY 206
Db 852 VYTLPPSRDELTKNQVSLTCLVKGYPSDIAVEWESNGQPENNYKTTTPVLDSDGSFFLY 911
Qy 207 SKLTVDKSRWQGNVFSQVSMHEALHNNHYTKLSLSPGK 246
Db 912 SKLTVDKSRWQGNVFSQVSMHEALHNNHYTKLSLSPGK 951
```

```
RESULT 4
US-09-935-868-7
; Sequence 7, Application US/09935868
; Patent No. US20020164690A1
; GENERAL INFORMATION:
; APPLICANT: Regeneron Pharmaceuticals, Inc
; TITLE OF INVENTION: Receptor Based Antagonists, and Methods of Making and Using
; FILE REFERENCE: REG 203D
; CURRENT APPLICATION NUMBER: US/09/935.868
; PRIOR FILING DATE: 2002-04-11
; PRIOR APPLICATION NUMBER: PCT/US99/22045
; PRIOR FILING DATE: 1999-09-22
; NUMBER OF SEQ ID NOS: 52
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 7
; LENGTH: 859
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-935-868-7

Query Match 94.0%; Score 1254; DB 9; Length 859;
Best Local Similarity 60.5%; Pred. No. 3.6e-32;
Matches 233; Conservative 8; Mismatches 5; Indels 139; Gaps 3;

Qy 1 DW-----LKAF-----6
Db 469 DWQOEDGTVHRTYLRGNLAESKCYLITVPVYADPGSPESIKAYLKQAPSGKGTVRTK 528
Qy 7 -----YDKVAEKLKEAFM-----19
Db 529 KVGKNEAVLEWDQLPVDVQNGFIRNYTIFYRTIIGNETAVNVDSSTHTYTLTSLTSDTLY 588
```

```
Qy 20 -----DKTHTCCPPCAPPELLGGPSVFL 41
Db 589 MVRMAAYTDEGGKDGPEFTFTPKFAQGEIESGEPKSCDKTHTCPAPPELLGGPSVFL 648
Qy 42 FPPKPKDTLMISRTPEVTCVVDVSHEDPEVKFNWYVDGVVEVHNAKTKPREEQYNSTYRV 101
Db 649 FPPKPKDTLMISRTPEVTCVVDVSHEDPEVKFNWYVDGVVEVHNAKTKPREEQYNSTYRV 708
Qy 102 VSVLTVLHODWLNKGEYCKVSKNKPALPAPIEKTISKAKGQPREPQVYTLPPSRDELTKNQ 161
Db 709 VSVLTVLHODWLNKGEYCKVSKNKPALPAPIEKTISKAKGQPREPQVYTLPPSRDELTKNQ 768
Qy 162 VSLCLVKGYPSDIAVEWESNGQPENNYKTTTPVLDSDGSFFLYSKLTVDKSRWQGNV 221
Db 769 VSLCLVKGYPSDIAVEWESNGQPENNYKTTTPVLDSDGSFFLYSKLTVDKSRWQGNV 828
Qy 222 FSCSVMHREALHNNHYTKLSLSPGK 246
Db 829 FSCSVMHREALHNNHYTKLSLSPGK 853

RESULT 5
US-09-935-868-26
; Sequence 26, Application US/09935868
; Patent No. US20020164690A1
; GENERAL INFORMATION:
; APPLICANT: Regeneron Pharmaceuticals, Inc
; TITLE OF INVENTION: Receptor Based Antagonists, and Methods of Making and Using
; FILE REFERENCE: REG 203D
; CURRENT APPLICATION NUMBER: US/09/935.868
; PRIOR FILING DATE: 2002-04-11
; PRIOR APPLICATION NUMBER: PCT/US99/22045
; PRIOR FILING DATE: 1999-09-22
; NUMBER OF SEQ ID NOS: 52
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 26
; LENGTH: 1158
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-935-868-26

Query Match 94.0%; Score 1254; DB 9; Length 1158;
Best Local Similarity 61.2%; Pred. No. 7.2e-32;
Matches 233; Conservative 8; Mismatches 5; Indels 135; Gaps 3;

Qy 1 DW-----LKAF-----6
Db 778 DWQOEDGTVHRTYLRGNLAESKCYLITVPVYADPGSPESIKAYLKQAPSGKGTVRTK 837
Qy 7 -----YDKVAEKLKEAFM-----19
Db 838 KVGKNEAVLEWDQLPVDVQNGFIRNYTIFYRTIIGNETAVNVDSSTHTYTLTSLTSDTLY 897
Qy 20 -----DKTHTCCPPCAPPELLGGPSVFLFPPK 45
Db 898 MVRMAAYTDEGGKDGPEFTFTPKFAQGEIESGEPKSCDKTHTCPAPPELLGGPSVFLFPPK 957
Qy 46 PKDTLMISRTPEVTCVVDVSHEDPEVKFNWYVDGVVEVHNAKTKPREEQYNSTYRVSVL 105
Db 958 PKDTLMISRTPEVTCVVDVSHEDPEVKFNWYVDGVVEVHNAKTKPREEQYNSTYRVSVL 1017
Qy 106 TVLHODWLNKGEYCKVSKNKPALPAPIEKTISKAKGQPREPQVYTLPPSRDELTKNQVSLT 165
Db 1018 TVLHODWLNKGEYCKVSKNKPALPAPIEKTISKAKGQPREPQVYTLPPSRDELTKNQVSLT 1077
Qy 166 CLVKGFPDIAVEWESNGQPENNYKTTTPVLDSDGSFFLYSKLTVDKSRWQGNVFS 225
Db 1078 CLVKGFPDIAVEWESNGQPENNYKTTTPVLDSDGSFFLYSKLTVDKSRWQGNVFS 1137
Qy 226 VMHEALHNNHYTKLSLSPGK 246
Db 1138 VMHEALHNNHYTKLSLSPGK 1158
```

```

RESULT 6
US-09-935-868-24
; Sequence 24, Application US/09935868
; Patent No. US20020164690A1
; GENERAL INFORMATION:
; APPLICANT: Regeneron Pharmaceuticals, Inc
; TITLE OF INVENTION: Receptor Based Antagonists, and Methods of Making and Using
; FILE REFERENCE: REG 203D
; CURRENT APPLICATION NUMBER: US/09/935,868
; CURRENT FILING DATE: 2002-04-11
; PRIOR APPLICATION NUMBER: PCT/US99/22045
; PRIOR FILING DATE: 1999-09-22
; NUMBER OF SEQ ID NOS: 52
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 24
; LENGTH: 1168
; TYPE: PRP
; ORGANISM: Homo sapiens
US-09-935-868-24

Query Match 94.0%; Score 1254; DB 9; Length 1168;
Best Local Similarity 61.2%; Pred. No. 7.3e-32;
Matches 233; Conservative 8; Mismatches 5; Indels 135; Gaps 3;

QY 1 DW-----LKAF----- 6
||
Db 788 DMOQEDGTVHRTYLRGNLAESKCYLITVTPYADGPGSPESIKAYLKQAPPSKGPTVRTK 847
:||||:
QY 7 -----YDKVAEKLEAF----- 19
:||||:
Db 848 KVGKNEAVLEWDQLPVDVQNGFIRNTYFRTIIGNETAVNVDSHTEYTLSSLTSDTL 907
:||||:
QY 20 -----DKTHTCPPCPAPPELLGGPSVFLFPPK 45
|||||
Db 908 MVRMAAYTDGGKDGPEFTTTPKFAQGETESGDKTHTCPPCPAPPELLGGPSVFLFPPK 967
|||||
QY 46 PKDTLMISRTPEVTCVVDVSHEDPEVKFNWYVDGVEVHNATKPREQYNSTYRVVSVL 105
|||||
Db 968 PKDTLMISRTPEVTCVVDVSHEDPEVKFNWYVDGVEVHNATKPREQYNSTYRVVSVL 1027
|||||
QY 106 TVLHQDLNKGKEYCKVSKNKPAPTEKTSKAKGQPREPQVYTLPPSRDELTKNOVSLT 165
|||||
Db 1028 TVLHQDLNKGKEYCKVSKNKPAPTEKTSKAKGQPREPQVYTLPPSRDELTKNOVSLT 1087
|||||
QY 166 CLVKGFPSPDIAVWESNGOPENNYKTPPVLDSDGSFFLYSKLTVDKSRWQGNVFS 225
|||||
Db 1088 CLVKGFPSPDIAVWESNGOPENNYKTPPVLDSDGSFFLYSKLTVDKSRWQGNVFS 1147
|||||
QY 226 VMHEALHNHYTQKSLSPCK 246
|||||
Db 1148 VMHEALHNHYTQKSLSPCK 1168
|||||

RESULT 7
US-09-825-561A-16
; Sequence 16, Application US/09825561A
; Patent No. US20020137677A1
; GENERAL INFORMATION:
; APPLICANT: Sprecher, Cindy A.
; APPLICANT: No. US20020137677A1ak, Julia E.
; APPLICANT: West, James W.
; APPLICANT: Presnell, Scott R.
; APPLICANT: Holly, Richard D.
; APPLICANT: Nelson, Andrew J.
; TITLE OF INVENTION: SOLUBLE ZALPHA11 CYTOKINE RECEPTORS
; FILE REFERENCE: 00-22
; CURRENT APPLICATION NUMBER: US/09/825,561A
; CURRENT FILING DATE: 2000-04-05
; PRIOR APPLICATION NUMBER: US 60/194,731
; PRIOR FILING DATE: 2000-04-05
; PRIOR APPLICATION NUMBER: US 60/222,121
; PRIOR FILING DATE: 2000-07-28

; NUMBER OF SEQ ID NOS: 86
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 16
; LENGTH: 567
; TYPE: PRP
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: soluble zalphallR/IgGgamma1 polypeptide
US-09-825-561A-16

Query Match 93.9%; Score 1253; DB 10; Length 567;
Best Local Similarity 55.9%; Pred. No. 1.5e-32;
Matches 236; Conservative 3; Mismatches 3; Indels 180; Gaps 3;

QY 5 AFY----- 7
|||
Db 146 AFYMLGKLOYELOQYRNGDPWAVSPRKLISVDSRSVLLPLEFRKDSSEYELQVRGPM 205
:||||:
QY 8 -----DKV-----AEKLKEAF----- 18
|||
Db 206 PGSSYQGTWSESDPVIFOTQSEELKEGWNPHASTKGPSVFPLAPSSKSTSGGTAALGCL 265
:||||:
QY 19 ----- 18
Db 266 VKDYFPPEPTVSNWNGALTSGVHTFFAVLQSSGLYSLSSVTVVPSSSLGTQTYICNVNKH 325
:||||:
QY 19 -----MDKTHTCPPCPAPPELLGGPSVFLFPPKPKDTLMISRTPEVTCVVD 64
|||||
Db 326 PSNTKVDKVKVEPKSCDKTHTCPPCPAPPELLGGPSVFLFPPKPKDTLMISRTPEVTCVVD 365
|||||
QY 65 VSHEDPEVKFNWYVDGVEVHNATKPREQYNSTYRVVSVLTVLHQDLNKGKEYCKVSN 124
|||||
Db 386 VSHEDPEVKFNWYVDGVEVHNATKPREQYNSTYRVVSVLTVLHQDLNKGKEYCKVSN 445
|||||
QY 125 KALPAPIETKISKAKGQPREPQVYTLPPSRDELTKNOVSLTCLVKGFYPSDIAVWESNG 184
|||||
Db 446 KALPAPIETKISKAKGQPREPQVYTLPPSRDELTKNOVSLTCLVKGFYPSDIAVWESNG 505
|||||
QY 185 QPENNYKTPPVLDSDGSFFLYSKLTVDKSRWQGNVFSVMHEALHNHYTQKSLSP 244
|||||
Db 506 QPENNYKTPPVLDSDGSFFLYSKLTVDKSRWQGNVFSVMHEALHNHYTQKSLSP 565
|||||
QY 245 GK 246
||
Db 566 GK 567
|||

RESULT 8
US-09-740-002-27
; Sequence 27, Application US/09740002
; Patent No. US20020001798A1
; GENERAL INFORMATION:
; APPLICANT: BRAMS, PETER
; APPLICANT: MORROW, PHILLIP
; TITLE OF INVENTION: NEUTRALIZING HIGH AFFINITY HUMAN MONOCLONAL ANTIBODIES
; TITLE OF INVENTION: SPECIFIC TO RSV F-PROTEIN AND METHODS FOR THEIR
; TITLE OF INVENTION: MANUFACTURE AND THERAPEUTIC USE THEREOF
; FILE REFERENCE: 037003-0275759
; CURRENT APPLICATION NUMBER: US/09/740,002
; CURRENT FILING DATE: 2000-12-20
; PRIOR APPLICATION NUMBER: 09/335,697
; PRIOR FILING DATE: 1999-06-18
; PRIOR APPLICATION NUMBER: 08/488,376
; PRIOR FILING DATE: 1995-06-07
; NUMBER OF SEQ ID NOS: 27
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 27
; LENGTH: 475
; TYPE: PRP
; ORGANISM: Homo sapiens
US-09-740-002-27

Query Match 93.8%; Score 1251; DB 10; Length 475;
```

Best Local Similarity 57.2%; Pred. No. 1.2e-32;  
Matches 234; Conservative 4; Mismatches 8; Indels 163; Gaps 2;

QY 1 DWL-----KAFYD--- 8  
Db 67 EWLARIDWDDTTFYSASLTKRLSLKSDTKSKNQVVLRTNVDPVDTATYFCARASLYDSDS 126  
QY 9 ----- 8  
Db 127 FYLFYHAYWGQTVTVSSASTKGPSVFPLAPSSKSTSGGTAALGCLVKDYFPEPVTVSW 186  
QY 9 -----KVAEKLKEA 17  
Db 187 NSGALTSVHTFPAVLQSSGLYSLSSVTVTPSSSLGTQTYICNVNHPKPSNTKVDKKAEPK 246  
QY 18 FMDKTHCPCPAPPELLGGPSVFLFPPPKDPTLMISRTPEVTCVVDVSHEDPEVFENWY 77  
Db 247 SCDKTHCPCPAPPELLGGPSVFLFPPPKDPTLMISRTPEVTCVVDVSHEDPEVFENWY 306  
QY 78 VDGVEVHNKTKPREQYNSTYRVSVTLVHODWLNKGKEYCKVSNKALPAPIETISK 137  
Db 307 VDGVEVHNKTKPREQYNSTYRVSVTLVHODWLNKGKEYCKVSNKALPAPIETISK 366  
QY 138 AKGQPREPQVYTLPPSRDELTKNQVSLTCLVKGFYPSDIAVEWESNGQPENNYKTTTPPV 197  
Db 367 AKGQPREPQVYTLPPSRDELTKNQVSLTCLVKGFYPSDIAVEWESNGQPENNYKTTTPPV 426  
QY 198 DSPGSEFLYSLKLVDSRWQGNVFCSCVMHEALHNNHYTKQSLSPGK 246  
Db 427 DSPGSEFLYSLKLVDSRWQGNVFCSCVMHEALHNNHYTKQSLSPGK 475

RESULT 9  
US-09-747-669-3  
; Sequence 3, Application US/09747669  
; Patent No. US20020122807A1  
; GENERAL INFORMATION:  
; APPLICANT: Dan, Michael D.  
; APPLICANT: Saleh, Mansoor  
; TITLE OF INVENTION: ANTIGEN BINDING FRAGMENTS, DESIGNATED  
; TITLE OF INVENTION: 485 THAT SPECIFICALLY DETECT CANCER CELLS, NUCLEOTIDES  
; TITLE OF INVENTION: ENCODING THE FRAGMENTS, AND USE THEREOF FOR THE PROPHYLAXIS  
; TITLE OF INVENTION: AND DETECTION OF CANCERS  
; FILE REFERENCE: 316082001001  
; CURRENT APPLICATION NUMBER: US/09/747,669  
; CURRENT FILING DATE: 2002-04-08  
; PRIOR APPLICATION NUMBER: US 09/111,286  
; PRIOR FILING DATE: 1998-07-07  
; NUMBER OF SEQ ID NOS: 7  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 3  
; LENGTH: 476  
; TYPE: PRT  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Synthetic construct  
US-09-747-669-3

Query Match 93.8%; Score 1251; DB 10; Length 476;  
Best Local Similarity 55.1%; Pred. No. 1.2e-32;  
Matches 233; Conservative 6; Mismatches 7; Indels 177; Gaps 2;

QY 1 DWLK-----AFY----- 4  
Db 54 NWYRQPGQGLEWGMWNPNSGKTGYAQKFGQGVTRMTNTSIRYAYNELSGLRSDTAVY 113  
QY 5 -----AFY----- 7  
Db 114 FCARNADNVEMAAIYHYGDMVWGQGTVTTVSSASTKGPSVFPLAPSSKSTSGGTAALGC 173  
QY 8 ----- 7  
Db 174 LVKDYPPEPVTVSWNSGALTSVHTFPAVLQSSGLYSLSSVTVTPSSSLGTQTYICNVNH 233

QY 8 ----DKVAEKLKEAFMDKTHCPCPAPPELLGGPSVFLFPPPKDPTLMISRTPEVTCV 63  
Db 234 KPSNTKVDKVEPKSCDKTHCPCPAPPELLGGPSVFLFPPPKDPTLMISRTPEVTCV 293  
QY 64 DVSHEDPEVKFNWYVDGVEVHNKTKPREQYNSTYRVSVTLVHODWLNKGKEYCKV 123  
Db 294 DVSHEDPEVKFNWYVDGVEVHNKTKPREQYNSTYRVSVTLVHODWLNKGKEYCKV 353  
QY 124 NKALPAPIETISKAKGQPREPQVYTLPPSRDELTKNQVSLTCLVKGFYPSDIAVEWESN 183  
Db 354 NKALPAPIETISKAKGQPREPQVYTLPPSRDELTKNQVSLTCLVKGFYPSDIAVEWESN 413  
QY 184 GQPENNYKTTTPPVLDSDGSEFLYSLKLVDSRWQGNVFCSCVMHEALHNNHYTKQSL 243  
Db 414 GQPENNYKTTTPPVLDSDGSEFLYSLKLVDSRWQGNVFCSCVMHEALHNNHYTKQSL 473  
QY 244 PGK 246  
Db 474 PGK 476

RESULT 10  
US-09-832-659-2  
; Sequence 2, Application US/09832659  
; Patent No. US20020155547A1  
; GENERAL INFORMATION:  
; APPLICANT: BIOGEN, INC.  
; TITLE OF INVENTION: Interferon-Beta Fusion Proteins and Uses  
; FILE REFERENCE: A064PCTSEQ  
; CURRENT APPLICATION NUMBER: US/09/832,659  
; CURRENT FILING DATE: 2001-04-11  
; PRIOR APPLICATION NUMBER: 60/120,237  
; PRIOR FILING DATE: 1999-02-16  
; PRIOR APPLICATION NUMBER: 60/104,491  
; PRIOR FILING DATE: 1998-10-16  
; NUMBER OF SEQ ID NOS: 44  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 2  
; LENGTH: 399  
; TYPE: PRT  
; ORGANISM: murine  
US-09-832-659-2

Query Match 93.7%; Score 1250; DB 9; Length 399;  
Best Local Similarity 83.5%; Pred. No. 8.6e-33;  
Matches 232; Conservative 5; Mismatches 7; Indels 34; Gaps 1;

QY 3 LKAFYDKVAEKLK-----EAFMDKTHTCPPC 28  
Db 122 LKRYGGRILHYLKAKEYSHCAWTVRVEILRNFYFINRLTGYLRNDDDKVDKTHTCPPC 181  
QY 29 PAPELLGGPSVFLFPPPKDPTLMISRTPEVTCVVDVSHEDPEVKFNWYVDGVEVHNK 88  
Db 182 PAPELLGGPSVFLFPPPKDPTLMISRTPEVTCVVDVSHEDPEVKFNWYVDGVEVHNK 241  
QY 89 KPREQYNSTYRVSVTLVHODWLNKGKEYCKVSNKALPAPIETISKAKGQPREPQV 148  
Db 242 KPREQYNSTYRVSVTLVHODWLNKGKEYCKVSNKALPAPIETISKAKGQPREPQV 301  
QY 149 TLPPSRDELTKNQVSLTCLVKGFYPSDIAVEWESNGQPENNYKTTTPPVLDSDGSEFLY 208  
Db 302 TLPPSRDELTKNQVSLTCLVKGFYPSDIAVEWESNGQPENNYKTTTPPVLDSDGSEFLY 361  
QY 209 LTVDSRWQGNVFCSCVMHEALHNNHYTKQSLSPGK 246  
Db 362 LTVDSRWQGNVFCSCVMHEALHNNHYTKQSLSPGK 399

RESULT 11  
US-09-832-659-44  
; Sequence 44, Application US/09832659  
; Patent No. US20020155547A1

GENERAL INFORMATION:  
 APPLICANT: BIOGEN, INC.  
 TITLE OF INVENTION: Interferon-Beta Fusion Proteins and Uses  
 FILE REFERENCE: A064PCTSEQ  
 CURRENT APPLICATION NUMBER: US/09/832,659  
 CURRENT FILING DATE: 2001-04-11  
 PRIOR APPLICATION NUMBER: 60/120,237  
 PRIOR FILING DATE: 1999-02-16  
 PRIOR APPLICATION NUMBER: 60/104,491  
 PRIOR FILING DATE: 1998-10-16  
 NUMBER OF SEQ ID NOS: 44  
 SOFTWARE: PatentIn Ver. 2.0  
 SEQ ID NO 44

TYPE: PRT  
 LENGTH: 423  
 ORGANISM: Homo sapiens  
 US-09-832-659-44

Query Match 93.7%; Score 1250; DB 9; Length 423;  
 Best Local Similarity 83.5%; Pred. No. 9.8e-33;  
 Matches 232; Conservative 4; Mismatches 8; Indels 34; Gaps 1;

QY 3 LKAFYDKVAEKLKE-----AFMDKTHTCPPC 28  
 |||||  
 DB 146 LKRYGRILHLKAEKESHCATIVRVEILRNIFYINRLTCVLRNGGSGVDKTHTCPPC 205  
 QY 29 PAPELLGGPSVFLPPKPKDTLMISRTPEVTCVVVDVSHEDPEVKFNWYVDGVEVHNAKT 88  
 |||||  
 DB 206 PAPELLGGPSVFLPPKPKDTLMISRTPEVTCVVVDVSHEDPEVKFNWYVDGVEVHNAKT 265  
 QY 89 KPREEQYNSTYRVVSVLTVLHQDWLNGKEYKCKVSNKALPAPIETISKAKGQPREPQVY 148  
 |||||  
 DB 266 KPREEQYNSTYRVVSVLTVLHQDWLNGKEYKCKVSNKALPAPIETISKAKGQPREPQVY 325  
 QY 149 TLPPSRDELTKNQVSLTCLVKGFYPSDIAVEWESNGQPENNYKTTPPVLDSDGSFFLYSK 208  
 |||||  
 DB 326 TLPPSRDELTKNQVSLTCLVKGFYPSDIAVEWESNGQPENNYKTTPPVLDSDGSFFLYSK 385  
 QY 209 LTVDKSRWQOGNVSFCSVMHEALHNHYTQKSLSLSPGK 246  
 |||||  
 DB 386 LTVDKSRWQOGNVSFCSVMHEALHNHYTQKSLSLSPGK 423

RESULT 12  
 US-10-091-236-17  
 Sequence 17, Application US/10091236  
 Patent No. US20020168360A1  
 GENERAL INFORMATION:

APPLICANT: DINGIVAN, CHRISTINE A.  
 TITLE OF INVENTION: METHODS OF PREVENTING OR TREATING INFLAMMATORY OR AUTOIMMUNE  
 DISORDERS BY ADMINISTERING INTEGRIN ALPHA-V-BETA-3 ANTAGONISTS  
 TITLE OF INVENTION: COMBINATION WITH OTHER PROPHYLACTIC OR THERAPEUTIC AGENTS  
 FILE REFERENCE: 10271-053-999  
 CURRENT APPLICATION NUMBER: US/10/091,236  
 CURRENT FILING DATE: 2002-03-04  
 PRIOR APPLICATION NUMBER: US 60/273,098  
 PRIOR FILING DATE: 2001-03-02  
 PRIOR APPLICATION NUMBER: US 60/316,321  
 PRIOR FILING DATE: 2001-08-31  
 NUMBER OF SEQ ID NOS: 17  
 SOFTWARE: PatentIn version 3.1  
 SEQ ID NO 17  
 LENGTH: 347  
 TYPE: PRT  
 ORGANISM: Homo sapiens  
 US-10-091-236-17

Query Match 93.6%; Score 1249; DB 9; Length 347;  
 Best Local Similarity 80.2%; Pred. No. 6.7e-33;  
 Matches 235; Conservative 1; Mismatches 8; Indels 49; Gaps 2;

QY 2 WLKAFYDKVAEKLKEAF----- 18

DB 56 WKQK-KDKVAELENSEFRASFKNRVYLDTVSGSLTYINTLTSSDEDEYEMESPNTDTM 114  
 QY 19 -----MDKTHTCPPCPAPELGGPSVFLPPKPKDTLMISRTPEVTCVVVDVSHEDPEVK 73  
 |||||  
 DB 115 KFFLYVDKTHTCPPCPAPELGGPSVFLPPKPKDTLMISRTPEVTCVVVDVSHEDPEVK 174  
 QY 74 FNWYVDGVEVHNAKTKPREEQYNSTYRVVSVLTVLHQDWLNGKEYKCKVSNKALPAPIEK 133  
 |||||  
 DB 175 FNWYVDGVEVHNAKTKPREEQYNSTYRVVSVLTVLHQDWLNGKEYKCKVSNKALPAPIEK 234  
 QY 134 TISKAKGQPREPQVYTLPPSRDELTKNQVSLTCLVKGFYPSDIAVEWESNGQPENNYKT 193  
 |||||  
 DB 235 TISKAKGQPREPQVYTLPPSRDELTKNQVSLTCLVKGFYPSDIAVEWESNGQPENNYKT 294  
 QY 194 PPVLDSDGSFFLYSKLTVDKSRWQOGNVSFCSVMHEALHNHYTQKSLSLSPGK 246  
 |||||  
 DB 295 PPVLDSDGSFFLYSKLTVDKSRWQOGNVSFCSVMHEALHNHYTQKSLSLSPGK 347

RESULT 13

US-10-091-313-7  
 Sequence 7, Application US/10091313  
 Publication No. US20030044406A1  
 GENERAL INFORMATION:

APPLICANT: DINGIVAN, CHRISTINE  
 TITLE OF INVENTION: METHODS OF PREVENTING OR TREATING INFLAMMATORY OR AUTOIMMUNE  
 DISORDERS BY ADMINISTERING CD2 ANTAGONISTS IN COMBINATION WITH  
 TITLE OF INVENTION: PROPHYLACTIC OR THERAPEUTIC AGENTS  
 FILE REFERENCE: 10271-063  
 CURRENT APPLICATION NUMBER: US/10/091,313  
 CURRENT FILING DATE: 2002-03-04  
 PRIOR APPLICATION NUMBER: US 60/273,098  
 PRIOR FILING DATE: 2001-03-02  
 PRIOR APPLICATION NUMBER: US 60/346,918  
 PRIOR FILING DATE: 2001-10-19  
 NUMBER OF SEQ ID NOS: 13  
 SOFTWARE: PatentIn version 3.1  
 SEQ ID NO 7  
 LENGTH: 347  
 TYPE: PRT  
 ORGANISM: Homo Sapiens  
 US-10-091-313-7

Query Match 93.6%; Score 1249; DB 9; Length 347;  
 Best Local Similarity 80.2%; Pred. No. 6.7e-33;  
 Matches 235; Conservative 1; Mismatches 8; Indels 49; Gaps 2;

QY 2 WLKAFYDKVAEKLKEAF----- 18  
 |||||  
 DB 56 WKQK-KDKVAELENSEFRASFKNRVYLDTVSGSLTYINTLTSSDEDEYEMESPNTDTM 114  
 QY 19 -----MDKTHTCPPCPAPELGGPSVFLPPKPKDTLMISRTPEVTCVVVDVSHEDPEVK 73  
 |||||  
 DB 115 KFFLYVDKTHTCPPCPAPELGGPSVFLPPKPKDTLMISRTPEVTCVVVDVSHEDPEVK 174  
 QY 74 FNWYVDGVEVHNAKTKPREEQYNSTYRVVSVLTVLHQDWLNGKEYKCKVSNKALPAPIEK 133  
 |||||  
 DB 175 FNWYVDGVEVHNAKTKPREEQYNSTYRVVSVLTVLHQDWLNGKEYKCKVSNKALPAPIEK 234  
 QY 134 TISKAKGQPREPQVYTLPPSRDELTKNQVSLTCLVKGFYPSDIAVEWESNGQPENNYKT 193  
 |||||  
 DB 235 TISKAKGQPREPQVYTLPPSRDELTKNQVSLTCLVKGFYPSDIAVEWESNGQPENNYKT 294  
 QY 194 PPVLDSDGSFFLYSKLTVDKSRWQOGNVSFCSVMHEALHNHYTQKSLSLSPGK 246  
 |||||  
 DB 295 PPVLDSDGSFFLYSKLTVDKSRWQOGNVSFCSVMHEALHNHYTQKSLSLSPGK 347

RESULT 14

US-10-091-268-7  
 Sequence 7, Application US/10091268  
 Publication No. US20030068320A1  
 GENERAL INFORMATION:

APPLICANT: DINGIVAN, CHRISTINE A  
TITLE OF INVENTION: METHODS OF ADMINISTERING/DOSING CD2 ANTAGONISTS FOR THE PREVENTION OF  
FILE REFERENCE: 10271-054-999  
CURRENT APPLICATION NUMBER: US/10/091,268  
CURRENT FILING DATE: 2002-03-04  
PRIOR APPLICATION NUMBER: US 60/273,098  
PRIOR FILING DATE: 2001-03-02  
PRIOR APPLICATION NUMBER: US 60/346,918  
PRIOR FILING DATE: 2001-10-19  
NUMBER OF SEQ ID NOS: 7  
SOFTWARE: Patent in version 3.1  
SEQ ID NO 7  
LENGTH: 347  
TYPE: PRT  
ORGANISM: Homo sapiens  
US-10-091-268-7

Query Match 93.6%; Score 1249; DB 9; Length 347;  
Best Local Similarity 80.2%; Pred. No. 6.7e-33;  
Matches 235; Conservative 1; Mismatches 8; Indels 49; Gaps 2;  
QY 2 WLKAFYDKVAEKLKEAF----- 18  
DB 56 WKQK-KDKVAELENSEFRASFKNRVYLDTVSGSLTIYNLTSSDEDEYEMESPNTIDTM 114  
QY 19 -----MDKTHTCPPCAPPELLGGPSVFLPPPKPDTLMISRTPEVTCVVVDVSHEDPEVK 73  
DB 115 KFFLYVDKTHTCPPCAPPELLGGPSVFLPPPKPDTLMISRTPEVTCVVVDVSHEDPEVK 174  
QY 74 FNNYVDGVEVHNKTKPREEQYNSTYRVVSVLTVLHQDWLNGKEYKCKVSNKALPAPIEK 133  
DB 175 FNNYVDGVEVHNKTKPREEQYNSTYRVVSVLTVLHQDWLNGKEYKCKVSNKALPAPIEK 234  
QY 134 TISKAGQPREPQVYTLPPSRDELTKNQVSLTCLVKGFYPSDIAVWESNGQPENNYKT 193  
DB 235 TISKAGQPREPQVYTLPPSRDELTKNQVSLTCLVKGFYPSDIAVWESNGQPENNYKT 294  
QY 194 PPVLSDSGSFFLYSKLTVDKSRWQGNVFCSCVMHEALHNHYTQKSLSLSPGK 246  
DB 295 PPVLSDSGSFFLYSKLTVDKSRWQGNVFCSCVMHEALHNHYTQKSLSLSPGK 347

RESULT 15  
US-09-796-033-8  
Sequence 8, Application US/09796033  
Patent No. US20020009446A1  
GENERAL INFORMATION:  
APPLICANT: Magilavy, Daniel  
TITLE OF INVENTION: METHOD OF MODULATING MEMORY EFFECTOR  
FILE REFERENCE: 10274-044001  
CURRENT APPLICATION NUMBER: US/09/796,033  
CURRENT FILING DATE: 2001-02-27  
PRIOR APPLICATION NUMBER: PCT/US99/20026  
PRIOR FILING DATE: 1999-08-31  
PRIOR APPLICATION NUMBER: US 60/098,456  
PRIOR FILING DATE: 1998-08-31  
NUMBER OF SEQ ID NOS: 8  
SOFTWARE: FastSeq for Windows Version 4.0  
SEQ ID NO 8  
LENGTH: 347  
TYPE: PRT  
ORGANISM: Homo sapiens  
FEATURE:  
NAME/KEY: SIGNAL  
LOCATION: (1)...(28)  
US-09-796-033-8

Query Match 93.6%; Score 1249; DB 10; Length 347;  
Best Local Similarity 80.2%; Pred. No. 6.7e-33;  
Matches 235; Conservative 1; Mismatches 8; Indels 49; Gaps 2;

QY 2 WLKAFYDKVAEKLKEAF----- 18  
DB 56 WKQK-KDKVAELENSEFRASFKNRVYLDTVSGSLTIYNLTSSDEDEYEMESPNTIDTM 114  
QY 19 -----MDKTHTCPPCAPPELLGGPSVFLPPPKPDTLMISRTPEVTCVVVDVSHEDPEVK 73  
DB 115 KFFLYVDKTHTCPPCAPPELLGGPSVFLPPPKPDTLMISRTPEVTCVVVDVSHEDPEVK 174  
QY 74 FNNYVDGVEVHNKTKPREEQYNSTYRVVSVLTVLHQDWLNGKEYKCKVSNKALPAPIEK 133  
DB 175 FNNYVDGVEVHNKTKPREEQYNSTYRVVSVLTVLHQDWLNGKEYKCKVSNKALPAPIEK 234  
QY 134 TISKAGQPREPQVYTLPPSRDELTKNQVSLTCLVKGFYPSDIAVWESNGQPENNYKT 193  
DB 235 TISKAGQPREPQVYTLPPSRDELTKNQVSLTCLVKGFYPSDIAVWESNGQPENNYKT 294  
QY 194 PPVLSDSGSFFLYSKLTVDKSRWQGNVFCSCVMHEALHNHYTQKSLSLSPGK 246  
DB 295 PPVLSDSGSFFLYSKLTVDKSRWQGNVFCSCVMHEALHNHYTQKSLSLSPGK 347  
Search completed: April 21, 2003, 10:51:17  
Job time : 13.3353 secs





GenCore version 5.1.4.p5.4578  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: April 21, 2003, 10:42:26 ; Search time 18.0882 Seconds  
(without alignments)  
1307.428 Million cell updates/sec

Title: 7LINK2

Perfect score: 1334

Sequence: 1 DWLKAFYDKVAEKLKEAFMD.....MHEALHNHYTKLSLSLSPGK 246

Scoring table: BLOSUM62

Gapop 10.0 ; Gapext 0.0

Searched: 283224 seqs, 96134422 residues

Total number of hits satisfying chosen parameters: 283224

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

PIR\_73.\*

1: pir1.\*

2: pir2.\*

3: pir3.\*

4: pir4.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1243	93.2	330	1 GHU	Ig gamma-1 chain C
2	1234	92.5	374	2 S69339	Ig heavy chain V r
3	1233	92.4	255	4 S31866	Ig gamma-1 chain C
4	1190	89.2	234	2 PT0207	Ig gamma chain C r
5	1165	87.3	377	2 A23511	Ig gamma-3 chain C
6	1163	87.2	377	2 A60764	Ig gamma-3 chain C
7	1148	86.1	326	1 G2HU	Ig gamma-2 chain C
8	1144	85.8	327	1 G4HU	Ig gamma-4 chain C
9	1143	85.7	289	1 G3HUM	Ig gamma-4 chain C
10	923	69.2	323	1 GHRB	Ig gamma-3 heavy c
11	917	68.7	328	2 I47160	Ig gamma chain C r
12	917	68.7	328	2 I47159	Ig gamma 2a chain
13	911	68.3	277	2 I47162	Ig gamma 4 chain c
14	899	67.4	329	1 G2GP	Ig gamma-2 chain C
15	889	66.6	328	2 I47158	Ig gamma-1 chain c
16	885	66.3	328	2 I47161	Ig gamma 3 chain c
17	868	65.1	470	2 S22080	Ig heavy chain pre
18	866	64.9	472	2 S31459	Ig gamma-1 chain -
19	854	64.0	308	2 C30554	Ig heavy chain C r
20	853	63.9	329	1 G3MSC	Ig gamma-3 chain C
21	849	63.6	333	2 PS0018	Ig gamma-2b chain
22	847	63.5	444	2 PC4436	monoclonal antibod
23	842	63.1	398	1 G3MSM	Ig gamma-3 chain C
24	826	61.9	326	2 PS0017	Ig gamma-1 chain C
25	825	61.8	469	2 S37483	Ig gamma-2a chain
26	823	61.7	324	1 GIMS	Ig gamma-1 chain C
27	818	61.3	329	2 S00847	Ig gamma-2c chain
28	818	61.3	393	1 G1MSM	Ig gamma-1 chain C
29	814	61.0	330	1 G2MSA	Ig gamma-2a chain

#### ALIGNMENTS

##### RESULT 1

GHU

Ig gamma-1 chain C region - human

C:Species: Homo sapiens (man)

C:Date: 31-Jan-1981 #sequence.Revision 18-Aug-1982 #text.Change 16-Jul-1999

C:Accession: A93433; S36861; S36887; B90563; A90564; B91668; A91723; A02146

R:Ellison, J.W.; Berson, B.J.; Hood, L.E.

Nucleic Acids Res. 10, 4071-4079, 1982

A:Title: The nucleotide sequence of a human immunoglobulin C-gamma1 gene.

A:Reference number: A93433; MUID:82274238; PMID:6287432

A:Accession: A93433

A:Molecule type: DNA

A:Residues: 1-330 <ELL>

A:Cross-references: EMBL:Z17370

A:Note: This sequence has the G1m(17) allotypic marker, 97-Lys, and the G1m(1) marker

A:Note: Lys-330 is removed after translation

R:Harris, L.J.

Submitted to the EMBL Data Library, October 1992

A:Reference number: S33904

A:Accession: S36861

A:Molecule type: DNA

A:Residues: 2-330 <HAR>

A:Cross-references: EMBL:Z17370

R:Takahashi, N.; Ueda, S.; Obata, M.; Nikaido, T.; Nakai, S.; Honjo, T.

Cell 29, 671-679, 1982

A:Title: Structure of human immunoglobulin gamma genes: implications for evolution o

A:Reference number: S33887; MUID:83001943; PMID:6811139

A:Accession: S33887

A:Molecule type: DNA

A:Residues: 88-113;235-330 <TAK>

A:Cross-references: EMBL:Z17370

R:Cunningham, B.A.; Rutishauser, U.; Gall, W.E.; Gottlieb, P.D.; Waxdal, M.J.; Edelm

Biochemistry 9, 3161-3170, 1970

A:Title: The covalent structure of a human gammaG-immunoglobulin. VII. Amino acid se

A:Reference number: A90563; MUID:71064024; PMID:5489771

A:Contents: myeloma protein Eu

A:Accession: B90563

A:Molecule type: protein

A:Residues: 1-96,'R',98-135 <CUN>

A:Note: This sequence has the G1m(3) marker, 97-Arg

R:Rutishauser, U.; Cunningham, B.A.; Bennett, C.; Konigsberg, W.H.; Edelman, G.M.

Biochemistry 9, 3171-3181, 1970

A:Title: The covalent structure of a human gammaG-immunoglobulin. VIII. Amino acid s

A:Reference number: A90564; MUID:71064025; PMID:5530842

A:Contents: Eu

A:Accession: A90564

A:Molecule type: protein

A:Residues: 136-154,'Q',156-165,'Q',167-176,'Q',178-194,'N',196-197,'D',199-238,'E',

A:Note: this sequence has the G1m(non-1) markers, 239-Glu and 241-Met

R:Notstingl, H.; Hilschmann, N.

Hoppe-Sevler's Z. Physiol. Chem. 357, 1571-1604, 1976

A:Title: Die Primerstruktur eines monoklonalen IgG1-Immunglobulins (Myelomprotein N

```

igen Primaerstruktur.
A:Reference number: A91668; MUID:77070269; PMID:826475
A:Contents: myeloma protein Nie
A:Accession: B91668
A:Molecule type: protein
A:Residues: 1-34, 'Q', '36-96', 'K', '98-115', 'Q', '117-197', 'D', '199-238', 'D', '240', 'L', '242-268', 'E', '27
R:Schmidt, W.E.; Jung, H.D.; Palm, W.; Hilschmann, N.
Hoppe-Seyler's Z. Physiol. Chem. 364, 713-747, 1983
A:Title: Die Primaerstruktur des kristallisierbaren monoklonalen Immunglobulins IgG1 KOL
A:Reference number: A91723; MUID:83289131; PMID:6884994
A:Contents: myeloma protein KOL; disulfide bonds
A:Accession: A91723
A:Molecule type: protein
A:Residues: 1-96, 'R', '98-197', 'D', '199-238', 'E', '240', 'M', '242-266', 'D', '268-271', 'D', '273-330' <SCH
R:Gall, W.E.; Edelman, G.M.
Biochemistry 9, 3188-3196, 1970
A:Title: The covalent structure of a human gammaG-immunoglobulin. X. Intrachain disulfid
A:Reference number: A90565; MUID:71064027; PMID:4923144
A:Contents: annotation; disulfide bonds
R:Dreker, L.; Schwarz, J.; Reichel, W.; Hilschmann, N.
Hoppe-Seyler's Z. Physiol. Chem. 357, 1515-1540, 1976
A:Title: Rule of antibody structure. The primary structure of monoclonal IgG1 immunoglob
enbromide cleavage products, and the disulfide bridges.
A:Reference number: A91667; MUID:77070267; PMID:1002129
A:Contents: annotation; disulfide bonds
C:Genetics:
A:Gene: GDB:IGHG1
A:Cross-references: GDB:120085; OMIM:147100
A:Map position: 14q32.33-14q32.33
A:Introns: 99/1; 114/1; 224/1
C:Complex: An immunoglobulin heterotetramer subunit consists of two identical light (kap
hain disulfide bonds. In some cases, such as IgA and IgM, the subunits associate into la
C:Superfamily: immunoglobulin C region; immunoglobulin homology
F:20-85/Domain: immunoglobulin homology <IM1>
F:137-206/Domain: immunoglobulin homology <IM2>
F:243-310/Domain: immunoglobulin homology <IM3>
F:27-83,144-204,250-308/Disulfide bonds: #status experimental
F:103/Disulfide bonds: interchain (to light chain) #status experimental
F:109,112/Disulfide bonds: interchain (to heavy chain) #status experimental
F:180/Binding site: carbohydrate (Asn) (covalent) #status experimental

Query Match          93.2%; Score 1243; DB 1; Length 330;
Best Local Similarity 96.6%; Pred. No. 5.3e-35;
Matches 230; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

QY 9 KVAEKLKEAFMDKTHCPCPAPPELLGGPSVFLFPPKPKDTLMISRTPEVTCVYVDVSH 68
|||:::|||||
DB 93 KVDKVEPKSCDKTHCPCPAPPELLGGPSVFLFPPKPKDTLMISRTPEVTCVYVDVSH 152
QY 69 DPEVKFNWYVDGVEVHNATKPREQYNSTYRVVSVLTVLHQDWLNGKEYCKVSNKALP 128
|||||
DB 153 DPEVKFNWYVDGVEVHNATKPREQYNSTYRVVSVLTVLHQDWLNGKEYCKVSNKALP 212
QY 129 APIETKISKAGQPREPOVYTLPPSRDELTKNQVSLTCLVKGFYPSDIAVEWESNGQPEN 188
|||||
DB 213 APIETKISKAGQPREPOVYTLPPSRDELTKNQVSLTCLVKGFYPSDIAVEWESNGQPEN 272
QY 189 NYKTTTPVLDSDGSFFLYSLKLTVDKSRWQQGNVFSCSVMHEALHNHYTQKSLSLSPGK 246
|||||
DB 273 NYKTTTPVLDSDGSFFLYSLKLTVDKSRWQQGNVFSCSVMHEALHNHYTQKSLSLSPGK 330

RESULT 2
S69339
Ig heavy chain V region precursor - human
C:Species: Homo sapiens (man)
C:Date: 19-Mar-1997 #sequence_revision 19-Mar-1997 #text_change 01-Dec-2000
C:Accession: S69339; S72664
R:Khamilichi, A.A.; Aucouturier, P.; Preud'homme, J.L.; Cogne, M.
Eur. J. Biochem. 229, 54-60, 1995

```

```

A:Title: Structure of abnormal heavy chains in human heavy-chain-deposition disease
A:Reference number: S69339; MUID:95262687; PMID:7744049
A:Accession: S69339
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-374 <KHA>
A:Cross-references: EMBL:X81695
R:Khamilichi, A.A.
submitted to the EMBL Data Library, September 1994
A:Reference number: S72664
A:Accession: S72664
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-140, 'C', '142-374' <KH2>
A:Cross-references: EMBL:X81695
C:Superfamily: immunoglobulin C region; immunoglobulin homology

Query Match          92.5%; Score 1234; DB 2; Length 374;
Best Local Similarity 74.5%; Pred. No. 1.6e-34;
Matches 231; Conservative 6; Mismatches 7; Indels 66; Gaps 3;

QY 1 DWLKAFFY---DK----- 9
|||:::|||||
DB 67 EWLALIFWDDDDKRYSPSLRFLTITKDTSKNQVVLTMNVDPADTATYYCGYSVEGYGQG 126
QY 10 -----VAEKLEAFMDKTHCPCPAPPELLGGPSVFLFPPKPKDTLMISRT 56
|||:::|||||
DB 127 YRFHSMGQGLTVTVSSEPKSC--DKTHTCPCPAPPELLGGPSVFLFPPKPKDTLMISRT 184
QY 57 EYTCVVVDVSHEDPEVKFNWYVDGVEVHNATKPREQYNSTYRVVSVLTVLHQDWLNGK 116
|||||
DB 185 EVTCVVVDVSHEDPEVKFNWYVDGVEVHNATKPREQYNSTYRVVSVLTVLHQDWLNGK 244
QY 117 EYCKYVSNKALPAPIETKISKAGQPREPOVYTLPPSRDELTKNQVSLTCLVKGFYPSDI 176
|||||
DB 245 EYCKYVSNKALPAPIETKISKAGQPREPOVYTLPPSRDELTKNQVSLTCLVKGFYPSDI 304
QY 177 AVEWESNGQPENNYKTTTPVLDSDGSFFLYSLKLTVDKSRWQQGNVFSCSVMHEALHNHYT 236
|||||
DB 305 AVEWESNGQPENNYKTTTPVLDSDGSFFLYSLKLTVDKSRWQQGNVFSCSVMHEALHNHYT 364
QY 237 QKSLSLSPGK 246
|||||
DB 365 QKSLSLSPGK 374

```

```

RESULT 3
S31866
Ig gamma-1 chain C region - synthetic
C:Species: synthetic
A:Note: Homo sapiens (man) gene engineered and expressed in Escherichia coli
C:Date: 06-Jan-1995 #sequence_revision 17-Mar-1997 #text_change 19-May-2000
C:Accession: S31866
R:Filpula, D.
submitted to the EMBL Data Library, February 1993
A:Description: Screening method for protein-protein interactions of cloned gene produc
A:Reference number: S31866
A:Accession: S31866
A:Molecule type: mRNA
A:Residues: 1-255 <FIL>
A:Cross-references: EMBL:X70421; NID:g33068; PIDN:CAA49866.1; PID:g33069
C:Keywords: immunoglobulin
F:1-22/Region: Escherichia coli outer membrane protein A precursor
F:23-255/Region: human Ig gamma-1 chain C region

Query Match          92.4%; Score 1233; DB 4; Length 255;
Best Local Similarity 100.0%; Pred. No. 5e-35;
Matches 227; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 20 DKHTHTCPCPAPPELLGGPSVFLFPPKPKDTLMISRTPEVTCVYVDVSHEDPEVKFNWYVD 79
|||||
DB 29 DKHTHTCPCPAPPELLGGPSVFLFPPKPKDTLMISRTPEVTCVYVDVSHEDPEVKFNWYVD 88

```



Qy	18	FMDKTHT-----CPCCPAPELLGGSPVFLFPKPDKDTLMISR	54
		I : I : I :	
Db	76	FGTYTCNIDHDKPSNTKVDTVERKCVCPCPP- VAGSPVLFPPPKDKDTLMISR	134
Qy	55	TPEVTCTVDVSHEDPEVKFNWYDGVGEVINAKTKPREOYNSTYRVVSVLTVLHODWLN	114
Db	135	TPEVTCTVDVSHEDPEVQNWTVDGVEVINAKTKPREOFNSTFRVSVLTVVHODWLN	194
Qy	115	GKEYCKVSNKALPAPIEKTISKAKGPREPQVYTLPSSRDELTKNQVSLTCLVKGFYPS	174
Db	195	GKEYCKVSNKGLPAPIEKTISKTGQPREPQVYTLPSSREEMTKNQVSLTCLVKGFYPS	254
Qy	175	DIAVESNGQPENNYKTPPVLDSDGSFPLYSKLTVDKSRWQQGNVFSCSVMHEALHNH	234
Db	255	DIAVESNGQPENNYKTPPMILDSGSEFFLYSKLTVDKSRWQQGNVFSCSVMHEALHNH	314
Qy	235	YTKSLSLSPCK 246	
Db	315	YTKSLSLSPCK 326	
	RESULT 8		
	G4HU		
	Ig gamma-4 chain C region - human		
	C:Species: Homo sapiens (man)		
	C>Date: 02-Apr-1982 #sequence_revision 02-Apr-1982 #ext_change 16-Jul-1999		
	C:Accession: A90933; A90249; A02150		
	R:Elliison, J.; Buxbaum, J.; Hood, L.		
	DNA 1, 11-18, 1981		
	A:Title: Nucleotide sequence of a human immunoglobulin C-gamma4 gene.		
	A:Reference number: A90933; MUID:83157104; PMID:6299662		
	A:Accession: A90933		
	A:Molecule type: DNA		
	A:Residues: 1-327 <ELL>		
	A>Note: the sequence was determined from the germline gene		
	R:Pink, J.R.L.; Buttery, S.H.; De Vries, G.M.; Milstein, C.		
	Biochem. J. 117, 33-47, 1970		
	A:Title: Human immunoglobulin subclassess. Partial amino acid sequence of		
	A:Reference number: A90249; MUID:70207350; PMID:4192699		
	A:Accession: A90249		
	A:Molecule type: protein		
	A:Residues: 1-30;81-326 <PIN>		
	C:Genetics:		
	A:Gene: GDB:IGHG4		
	A:Cross-references: GDB:119340; OMIM:147130		
	A:Map position: 14q32.33-14q32.33		
	A:Citrons: 99/1; 111/1; 221/1		
	C:Complex: An immunoglobulin heterotetramer subunit consists of two identical disulfide bonds. In some cases, such as IgA and IgM, the subunits associate by interchain (to heavy chain) status predicted		
	C:Superfamily: immunoglobulin C region; immunoglobulin homology		
	F:Keywords: duplication; glycoprotein; heterotetramer; immunoglobulin		
	F:20-85/Domain: immunoglobulin homology <IM>		
	F:99-110/Region: hinge		
	F:134-203/Domain: immunoglobulin homology <IM2>		
	F:240-307/Domain: immunoglobulin homology <IM3>		
	F:14/Disulfide bonds: interchain (to light chain) #status experimental		
	F:27-83,141-201,247-305/Disulfide bonds: #status predicted		
	F:106,109/Disulfide bonds: interchain (to heavy chain) #status experimental		
	F:177/Binding site: carbohydrate (Asn) (covalent) #status predicted		
	Query Match 85.8%; Score 1144; DB 1; Length 327;		
	Best Local Similarity 84.7%; Pred. No. 1,3e-31;		
	Matches 211; Conservative 9; Mismatches 6; Indels 23; Gaps		
Qy	21	KTHT-----CPCCPAPELLGGSPVFLFPKPDKDTLMISR	57
		I : I :	
Db	79	KTYCWNVDHKPSNTKVDRKRYSKYGPCPCSPAPEPLGGSPVFLFPKPDKDTLMISR	138
Qy	58	VTCVTVDVSHEDPEVKFNWYDGVGEVINAKTKPREOYNSTYRVVSVLTVLHODWLNGKE	117
Db	139	VTCVTVDVSDQDPQVFNWYDGVGEVINAKTKPREOFNSTFRVSVLTVVHODWLNGKE	198

```

||||||| ||: |||||||||||||||||||||||||||||||||||||||
Db 199 YKCVSNKGLPSSTKTSKAKGPRQVYTLPPSQEEMTKNOVSLTCLVKGFPSDIA 258
QY 178 VEWESNGOPENNYKTTTPVLDSGDSFYLKSLTVDKSRWQGNVFSVMHEALHNHYTQ 237
||||||| ||: |||||||||||||||||||||||||||||||||||||||
Db 259 VEWESNGOPENNYKTTTPVLDSGDSFYLKSLTVDKSRWQGNVFSVMHEALHNHYTQ 318
QY 238 KSLSLSPGK 246
||||||| ||
Db 319 KSLSLGK 327

RESULT 9
G3HUW1
Ig gamma-3 heavy chain disease proteins - human
C:Species: Homo sapiens (man)
C:Date: 31-Dec-1979 #sequence_revision 23-Oct-1981 #text_change 16-Jul-1999
C:Accession: A90442; A92219; A90198; A93915; A02149
R:Frangione, B.; Rosenwasser, E.; Prelli, F.; Franklin, E.C.
Biochemistry 19, 4304-4308, 1980
A:Title: Primary structure of human gamma3 immunoglobulin deletion mutant: gamma3 heavy
A:Reference number: A90442; MUID:81021548; PMID:6774747
A:Contents: heavy chain disease protein wis
A:Accession: A90442
A:Molecule type: protein
A:Residues: 1-289 <FRA>
A:Note: the molecule is a dimer linked by 12 disulfide bonds; it has an extra interchain
A:Note: this protein lacks most of the V region and all of the CH1 region. Residue 12 co
A:Note: the sequence of residues 42-76 was taken from the reference that follows
R:Michaelson, T.E.; Frangione, B.; Franklin, E.C.
J. Biol. Chem. 252, 883-889, 1977
A:Title: Primary structure of the 'hinge' region of human IgG3. Probable quadruplication
A:Reference number: A92219; MUID:77118561; PMID:402363
A:Contents: normal gamma-3 chains, sequence corresponding to residues 12-97 of protein w
A:Accession: A92219
A:Molecule type: protein
A:Residues: 12-97 <MIC>
A:Note: the hinge region in gamma-3 chains is about four times as long as in other gamma
idue segment (12-28)
R:Wolfenstein-Rodei, C.; Frangione, B.; Prelli, F.; Franklin, E.C.
Biochem. Biophys. Res. Commun. 71, 907-914, 1976
A:Title: The amino acid sequence of "heavy chain disease" protein ZUC. Structure of the
A:Reference number: A90198; MUID:77021516; PMID:823945
A:Contents: heavy chain disease protein Zuc, partial sequence corresponding to residues
A:Accession: A90198
A:Molecule type: protein
A:Residues: 59-125, 'EB', 128-226, 228-289 <WOL>
A:Note: this protein lacks most of the V region, all of the CH1 region, and part of the
R:Alexander, A.; Steinmetz, M.; Barritault, D.; Frangione, B.; Franklin, E.C.; Hood, L.;
Proc. Natl. Acad. Sci. U.S.A. 79, 3260-3264, 1982
A:Title: gamma heavy chain disease in man: cDNA sequence supports partial gene deletion
A:Reference number: A93915; MUID:82247835; PMID:6808505
A:Contents: heavy chain disease protein Omn
A:Accession: A93915
A:Molecule type: mRNA
A:Residues: 12-70; 72-114; 116-125, 'E', 127-133, 'L', 135-136, 'E', 138, 'Y', 140-154, 'D', 156-157
A:Note: a carboxyl-terminal Lys is removed posttranslationally
A:Note: this sequence may represent an allelic form or another gamma chain subclass
C:Comment: The heavy chain disease protein wis is shown.
C:Genetics:
A:Gene: GDB:IGHG3
A:Cross-references: GDB:119339; OMIM:147120
A:Map position: 14q32.33-14q32.33
A:Superfamily: immunoglobulin C region; immunoglobulin homology
C:Keywords: duplication; glycoprotein; immunoglobulin; pyroglutamic acid
F:203-270/Domain: immunoglobulin homology <IMM>
F:1/Modified site: pyrrolidone carboxylic acid (Gln) #status experimental
F:6.140/Binding site: carbohydrate (Asn) (covalent) #status experimental

Query Match 85.7%; Score 1143; DB 1; Length 289;
Best Local Similarity 74.7%; Pred. No. 9.5e-32;
Matches 210; Conservative 15; Mismatches 11; Indels 45; Gaps 1;

```

```

QY 10 VAEKLEAFMDKTHT----- 24
||: |||
Db 9 VSSELKTPGLDTHTCRCPPEPKSCDTPPCRCPEPKSCDTPPCRCPEPKSCDTPPP 68
QY 25 CPCPAPPELLGGPSVFLFPKPKDITLMSRTEPTCVVVDVSHEDPEVKFNWYDGVGVH 84
||: |||
Db 69 CPCPAPPELLGGPSVFLFPKPKDITLMSRTEPTCVVVDVSHEDPEVKFNWYDGVGVH 128
QY 85 NAKTKPREQYNSTYRVVSVLTIVLHODWLNGKEYKCKVSNKALPAPIETKTISKAKGQPRE 144
||||||| ||: |||||||||||||||||||||||||||||||||||||||
Db 129 NAKTKPREQQFNSTFRVSVLTIVLHQNWLDSGKEYKCKVSNKALPAPIETKTISKAKGQPRE 188
QY 145 PQVYTLPPSRDELTKNOVSLTCLVKGFPSDIAVWESNGOPENNYKTTTPVLDSGDSFF 204
||||||| ||: |||||||||||||||||||||||||||||||||||||||
Db 189 PQVYTLPPSRDEMTKNOVSLTCLVKGFPSDIAVWESSGQPPENNYKTTTPMLDSGDSFF 248
QY 205 LYSKLTVDKSRWQGNVFSVMHEALHNHYTKLSLSPG 245
||||||| ||: |||||||||||||||||||||||||||||||||||||||
Db 249 LYSKLTVDKSRWQGNVFSVMHEALHNHYTKLSLSPG 289

RESULT 10
GHRB
Ig gamma chain C region - rabbit
C:Species: Oryctolagus cuniculus (domestic rabbit)
C:Date: 24-Apr-1984 #sequence_revision 15-Nov-1984 #text_change 16-Jul-1999
C:Accession: A91749; A90290; A93928; A90245; A94416; A02161
R:Bernstein, K.E.; Alexander, C.B.; Mage, R.G.
Immunogenetics 18, 387-397, 1983
A:Title: Nucleotide sequence of a rabbit IgG heavy chain from the recombinant F-I ha;
A:Reference number: A91749; MUID:84030930; PMID:6313520
A:Accession: A91749
A:Molecule type: mRNA
A:Residues: 1-323 <BER>
A:Note: this sequence has the d12 allotypic marker, 104-Thr, and the e14 marker, 185
R:Pratt, D.M.; Mole, L.E.
Biochem. J. 151, 337-349, 1975
A:Title: Sequence studies on the constant region of the Fd sections of rabbit immuno
A:Reference number: A90290; MUID:76135469; PMID:1243651
A:Accession: A90290
A:Molecule type: protein
A:Residues: 1-47, 'E', 49-71, 'PV', 72-128 <PRA>
R:Martens, C.L.; Moore, K.W.; Steinmetz, M.; Hood, L.; Knight, K.L.
Proc. Natl. Acad. Sci. U.S.A. 79, 6018-6022, 1982
A:Title: Heavy chain genes of rabbit IgG: isolation of a cDNA encoding gamma heavy c
A:Reference number: A93928; MUID:83299917; PMID:6193512
A:Accession: A93928
A:Molecule type: mRNA
A:Residues: 88-103, 'M', 105-143, 'E', 145-184, 'A', 186, 'E', 188-266 <MAR>
A:Cross-references: GB:M16426; NID:g165111; PIDN:AAA31289.1; PID:g165112
A:Note: this sequence has the d11 allotypic marker, 104-Met, and the e15 allotypic m
R:Fruchter, R.G.; Jackson, S.A.; Mole, L.E.; Porter, R.R.
Biochem. J. 116, 249-259, 1970
A:Title: Sequence studies of the Fd section of the heavy chain of rabbit immunoglobul
A:Reference number: A90245; MUID:70110015; PMID:5461106
A:Accession: A90245
A:Molecule type: protein
A:Residues: 132-143, 'E', 145-161 <FRU>
R:Hill, R.B.; Lebovitz, H.E.; Fellows Jr., R.E.; Delaney, R.
in Gamma Globulins, Nobel Symp. 3, Killander, J., ed., pp.109-127, Almqvist and Wiks
A:Reference number: A94416
A:Accession: A94416
A:Molecule type: protein
A:Residues: 129-131; 155-172, 'D', 174-184, 'A', 186, 'E', 188-200, 'D', 202-217, 'E', 219-232, '
A:Note: this has the e15 allotypic marker, 185-Ala
C:Complex: An immunoglobulin heterotetramer subunit consists of two identical light
hain disulfide bonds. In some cases, such as IgA and IgM, the subunits associate int
C:Superfamily: immunoglobulin C region; immunoglobulin homology
C:Keywords: duplication; glycoprotein; heterotetramer; immunoglobulin
F:20-82/Domain: immunoglobulin homology <IM1>
F:130-199/Domain: immunoglobulin homology <IM2>
F:236-303/Domain: immunoglobulin homology <IM3>

```









GenCore version 5.1.4.p5\_4578  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: April 21, 2003, 10:42:21 ; Search time 9.40588 Seconds  
(without alignments)  
1084.766 Million cell updates/sec

Title: 7LINK2  
Perfect score: 1334  
Sequence: 1 DWLKAFYDKVAELKEAFMD.....MHEALHNYTKSLSLSPGK 246

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.0

Searched: 112892 seqs, 41476328 residues

Total number of hits satisfying chosen parameters: 112892

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : SwissProt\_40:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	ID	Description
1	1243	93.2	GC1_HUMAN	P01857 homo sapien
2	1148	86.1	GC3_HUMAN	P01860 homo sapien
3	1148	86.1	GC2_HUMAN	P01859 homo sapien
4	1144	85.8	GC4_HUMAN	P01861 homo sapien
5	923	69.2	GC_RABIT	P01870 oryctolagus
6	899	67.4	GC2_CAVPO	P01862 cavia porce
7	853	63.9	GC3_MOUSE	P22436 mus musculus
8	849	63.6	GC3_MOUSE	P03987 mus musculus
9	842	63.1	GC3_MOUSE	P03987 mus musculus
10	826	61.9	GC1_RAT	P20759 rattus norv
11	823	61.7	GC1_MOUSE	P01868 mus musculus
12	818	61.3	GC1_MOUSE	P20762 rattus norv
13	818	61.3	GC1_MOUSE	P01869 mus musculus
14	814	61.0	GC3_MOUSE	P01863 mus musculus
15	809	60.6	GC3_MOUSE	P01864 mus musculus
16	809	60.6	GC3_MOUSE	P01865 mus musculus
17	795	59.6	GC3_MOUSE	P20760 rattus norv
18	789	59.1	GC3_MOUSE	P01866 mus musculus
19	784	58.8	GC3_MOUSE	P01867 mus musculus
20	388	29.1	EPC_MOUSE	P06336 mus musculus
21	381	28.6	GC1_MOUSE	P01871 homo sapien
22	378	28.3	EPC_MOUSE	P01855 rattus norv
23	377	28.3	GC1_MOUSE	P04220 homo sapien
24	377	28.3	GC1_MOUSE	P01872 mus musculus
25	376	28.2	GC1_MOUSE	P01854 homo sapien
26	373	28.0	GC1_MOUSE	P03988 oryctolagus
27	367	27.5	GC1_MOUSE	P01873 mus musculus
28	363	27.2	GC1_MOUSE	P04221 oryctolagus
29	356	26.7	GC1_MOUSE	P06337 mesocricetu
30	351	26.3	GC1_MOUSE	P20768 suncus muri
31	349	26.2	GC1_MOUSE	P01874 canis fami
32	334	25.0	HVC2_HETFR	P23085 heterodontu
33	332	24.9	HVC2_HETFR	P23087 heterodontu

34	326	24.4	446	1	MUC_CHICK	P01875 gallus gall
35	324	24.3	461	1	HVC2_HETFR	P23088 heterodontu
36	321	24.1	393	1	HVC3_HETFR	P23086 heterodontu
37	315	23.6	370	1	HVC1_HETFR	P23084 heterodontu
38	312	23.4	6669	1	NEBU_HUMAN	P20929 homo sapien
39	311	23.3	6486	1	TYCC_RACBR	O30409 b tyrocidin
40	310	23.2	299	1	ALC_RABIT	P01879 oryctolagus
41	308	23.1	6359	1	BAC2_BACLI	O68008 b bacitraci
42	306	22.9	4393	1	PGBM_HUMAN	P98160 homo sapien
43	303	22.7	4969	1	RYR2_RABIT	P30957 oryctolagus
44	302	22.6	353	1	ALC1_GORGO	P20758 gorilla gor
45	302	22.6	4829	1	BIR6_HUMAN	Q9nr09 homo sapien

ALIGNMENTS

RESULT 1	GC1_HUMAN	STANDARD;	PRT;	330 AA.
ID	GC1_HUMAN	STANDARD;	PRT;	330 AA.
AC	P01857:			
DT	21-JUL-1986 (Rel. 01, Created)			
DT	21-JUL-1986 (Rel. 01, Last sequence update)			
DT	16-OCT-2001 (Rel. 40, Last annotation update)			
DE	Ig gamma-1 chain C region.			
GN	IGHG1.			
OS	Homo sapiens (Human).			
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
OC	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.			
OX	NCBI_TaxID=9606;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RX	MEDLINE=82274238; PubMed=6287432;			
RA	Ellison J.W., Berson B.J., Hood L.E.;			
RT	"The nucleotide sequence of a human immunoglobulin C gamma1 gene."			
RL	Nucleic Acids Res. 10:4071-4079(1982).			
RN	[2]			
RP	SEQUENCE OF 1-135 (MYELOMA PROTEIN EU).			
RX	MEDLINE=71064024; PubMed=5489771;			
RA	Cunningham B.A., Rutishauser U., Gall W.E., Gottlieb P.D.,			
RT	Waxdal M.J., Edelman G.M.;			
RL	"The covalent structure of a human gamma G-immunoglobulin. VII. Amino acid sequence of heavy-chain cyanogen bromide fragments H1-H4."			
RN	[3]			
RP	Biochemistry 9:3161-3170(1970).			
RX	SEQUENCE OF 136-329 (EU).			
RA	MEDLINE=71064025; PubMed=5530842;			
RT	Rutishauser U., Cunningham B.A., Bennett C., Konigsberg W.H.,			
RL	Edelman G.M.;			
RT	"The covalent structure of a human gamma G-immunoglobulin. 8. Amino acid sequence of heavy-chain cyanogen bromide fragments H5-H7."			
RN	[4]			
RP	Biochemistry 9:3171-3181(1970).			
RX	SEQUENCE (MYELOMA PROTEIN NIE).			
RA	MEDLINE=77070269; PubMed=826475;			
RT	Ponstingl H., Hilschmann N.;			
RL	"The rule of antibody structure. The primary structure of a monoclonal IgG1 immunoglobulin (myeloma protein NIE). III. The chymotryptic peptides of the H-chain, alignment of the tryptic peptides and discussion of the complete structure."			
RN	[5]			
RP	Hoppe-Seyler's Z. Physiol. Chem. 357:1571-1604(1976).			
RX	SEQUENCE (MYELOMA PROTEIN KOL), AND DISULFIDE BONDS.			
RA	MEDLINE=83289131; PubMed=6884994;			
RT	Schmidt W.E., Jung H.-D., Palm W., Hilschmann N.;			
RL	"Three-dimensional structure determination of antibodies. Primary structure of crystallized monoclonal immunoglobulin IgG1 KOL, I."			
RN	[6]			
RP	Hoppe-Seyler's Z. Physiol. Chem. 364:713-747(1983).			
RX	DISULFIDE BONDS.			
RA	MEDLINE=71064027; PubMed=4923144;			
RT	Gall W.E., Edelman G.M.;			
RL	"The covalent structure of a human gamma G-immunoglobulin. X.			

RT Intrachain disulfide bonds.";  
 RL Biochemistry 9:3188-3196(1970).  
 [7]  
 RN DISULFIDE BONDS.  
 RX MEDLINE=77070267; PubMed=1002129;  
 RA Dreker L., Schwarz J., Reichel W., Hilschmann N.;  
 "Rule of antibody structure. The primary structure of a monoclonal  
 IgG1 immunoglobulin (myeloma protein Nie), I: Purification and  
 characterization of the protein, the L- and H-chains, the  
 cyanogen bromide cleavage products, and the disulfide bridges.";  
 RT Hoppe-Seyler's Z. Physiol. Chem. 357:1515-1540(1976).  
 [8]  
 RL  
 RN X-RAY CRYSTALLOGRAPHY (2.9 ANGSTROMS).  
 RP MEDLINE=81208100; PubMed=7236608;  
 RA Deisenhofer J.;  
 "Crystallographic refinement and atomic models of a human Fc fragment  
 and its complex with fragment B of protein A from Staphylococcus  
 aureus at 2.9- and 2.8-A resolution.";  
 RT Biochemistry 20:2361-2370(1981).  
 RL  
 CC GIM(1) MARKERS: NIE HAS THE GIM(17) ALLOTYPIC MARKER, 97-K, & THE  
 CC GIM(1) MARKERS, 239-D & 241-L. KOL & EU SEQUENCES HAVE THE GIM(3)  
 CC MARKER & THE GIM (NON-1) MARKERS.  
 CC -|- MISCELLANEOUS: NIE ALSO DIFFERS IN THE AMIDATION STATES OF  
 CC 35,116,198,269 & 272.  
 CC -|- MISCELLANEOUS: EU ALSO DIFFERS IN THE AMIDATION STATES OF RESIDUES  
 CC 155, 166, 177, 195, 198, 269, AND 272 AND IN THE ORDER OF RESIDUES  
 CC 268-272.  
 CC -|- MISCELLANEOUS: KOL ALSO DIFFERS IN THE AMIDATION STATES OF  
 CC RESIDUES 198, 267&272.  
 CC  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC  
 DR EMBL; J00228; AAC82527.1; ALT\_INIT.  
 DR PIR; A02146; GHHD.  
 DR PDB; 1FC1; 15-JUL-92.  
 DR PDB; 1FC2; 15-JUL-92.  
 DR Genew; HGNC:5525; IGHG1.  
 DR MIM; 147100; -.  
 DR InterPro; IPR003006; Ig\_MHC.  
 DR InterPro; IPR003597; Ig\_C1.  
 DR InterPro; IPR003600; Ig\_Like.  
 DR Pfam; PF00047; Ig; 3.  
 DR SMART; SM00410; Ig\_Like; 1.  
 DR SMART; SM00407; IGHG1; 2.  
 DR PROSITE; PS00290; Ig\_MHC; 2.  
 KW Immunoglobulin domain; Immunoglobulin C region; Glycoprotein;  
 KW 3D-structure.  
 FT NON\_TER 1 1  
 FT DOMAIN 1 98  
 FT DOMAIN 99 110  
 FT HINGE.  
 FT DOMAIN 111 223  
 FT CH2.  
 FT DOMAIN 224 330  
 FT CH3.  
 FT DISULFID 27 83  
 FT DISULFID 103 103  
 FT DISULFID 109 109  
 FT DISULFID 112 112  
 FT DISULFID 144 204  
 FT DISULFID 250 308  
 FT CARBOHYD 180 180  
 FT MOD\_RES 330 330  
 FT VARIANT 97 97  
 FT VARIANT 239 239  
 FT VARIANT 241 241  
 FT VARIANT 241 241  
 FT STRAND 123 126

FT HELIX 130 134  
 FT TURN 136 137  
 FT STRAND 141 148  
 FT STRAND 158 162  
 FT TURN 163 164  
 FT STRAND 165 166  
 FT STRAND 175 178  
 FT STRAND 183 190  
 FT HELIX 193 197  
 FT TURN 198 199  
 FT STRAND 202 206  
 FT STRAND 215 219  
 FT STRAND 227 227  
 FT STRAND 230 234  
 FT HELIX 238 240  
 FT TURN 241 242  
 FT STRAND 245 256  
 FT STRAND 260 266  
 FT TURN 267 268  
 FT STRAND 269 270  
 FT STRAND 274 276  
 FT STRAND 280 281  
 FT TURN 283 284  
 FT STRAND 287 296  
 FT HELIX 297 301  
 FT TURN 302 303  
 FT STRAND 306 312  
 FT TURN 313 314  
 FT TURN 316 317  
 FT STRAND 320 324  
 SQ SEQUENCE 330 AA; 36106 MW; 3770EE106C2FA33D CRC64;  
 Query Match 93.2%; Score 1243; DB 1; Length 330;  
 Best Local Similarity 96.6%; Pred. No. 1.6e-35;  
 Matches 230; Conservative 3; Mismatches 5; Indels 0; Gaps 0;  
 QY 9 KVAEKLKEAFMDKTHTCPPCAPPELLGGPSVFLPPKPKDTLMISRTPEVTCVVVDVSH 68  
 DB 93 KVDKKEPKSCDKTHTCTPCAPPELLGGPSVFLPPKPKDTLMISRTPEVTCVVVDVSH 152  
 QY 69 DPEVKFNWYVDGVEVHNAKTKPREEQYNSTYRVVSVLTVLHQDWLNGKEYKCKVSNKALP 128  
 DB 153 DPEVKFNWYVDGVEVHNAKTKPREEQYNSTYRVVSVLTVLHQDWLNGKEYKCKVSNKALP 212  
 QY 129 APIETISKAKGQPREPQVYTLPPSRDELTKNQVSLTCLVKGFYPSDIAVWESNGQPEN 188  
 DB 213 APIETISKAKGQPREPQVYTLPPSRDELTKNQVSLTCLVKGFYPSDIAVWESNGQPEN 272  
 QY 189 NYKTPPVLDSDGSFFLYSKLTVDKSRWQQGNVSCFVSYMSHALLNHYTQKSLSPGK 246  
 DB 273 NYKTPPVLDSDGSFFLYSKLTVDKSRWQQGNVSCFVSYMSHALLNHYTQKSLSPGK 330  
 RESULT 2  
 GC3\_HUMAN  
 ID GC3\_HUMAN STANDARD; PRT; 290 AA.  
 AC P01860;  
 DT 21-JUL-1986 (Rel. 01, Created)  
 DT 21-JUL-1986 (Rel. 01, Last sequence update)  
 DT 15-JUN-2002 (Rel. 41, Last annotation update)  
 DE Ig gamma-3 chain C region (Heavy chain disease protein) (HDC).  
 GN IGHG3.  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Euthera; Primates; Catarrhini; Hominiidae; Homo.  
 OX NCBI\_TaxID=9606;  
 RN [1]  
 RP SEQUENCE (DISEASE PROTEIN WIS).  
 RX MEDLINE=81021548; PubMed=6774747;  
 RA Frangione B., Rosenwasser E., Prelli F., Franklin E.C.;  
 "Primary structure of human gamma 3 immunoglobulin deletion mutant:  
 gamma 3 heavy-chain disease protein Wis.";  
 RT Biochemistry 19:4304-4308(1980).  
 RL

[2]  
 RN REVISIONS TO 12-97 (PROTEIN WIS).  
 RX MEDLINE=77118561; PubMed=402363;  
 RA Michaelson T.E., Frangione B., Franklin E.C.;  
 RT "Primary structure of the 'hinge' region of human IgG3. Probable  
 RL quadruplication of a 15-amino acid residue basic unit.";  
 RL J. Biol. Chem. 252:883-889(1977).  
 [3]  
 RN REVISIONS TO 59-289 (PROTEIN WIS/DISEASE PROTEIN ZUC).  
 RX MEDLINE=77021516; PubMed=823945;  
 RA Wolfenstein-Todel C., Frangione B., Prelli F., Franklin E.C.;  
 RT "The amino acid sequence of 'heavy chain disease' protein ZUC.  
 RL Structure of the Fc fragment of immunoglobulin G3.";  
 RL Biochem. Biophys. Res. Commun. 71:907-914(1976).  
 [4]  
 RN SEQUENCE FROM N.A. (DISEASE PROTEIN OMM).  
 RX MEDLINE=82247835; PubMed=6808505;  
 RA Alexander A., Steilmetz M., Barriault D., Frangione B.,  
 RA Franklin E.C., Hood L., Buxbaum J.N.;  
 RT "Gamma Heavy chain disease in man: cDNA sequence supports partial  
 RL gene deletion model.";  
 RL Proc. Natl. Acad. Sci. U.S.A. 79:3260-3264(1982).  
 CC -1- SUBUNIT: DIMER LINKED BY 12 DISULFIDE BONDS; IT HAS AN EXTRA  
 CC INTERCHAIN DISULFIDE BOND AT POSITION 7 IN ADDITION TO THE 11  
 CC NORMALLY PRESENT IN THE HINGE REGION.  
 CC -1- MISCELLANEOUS: THE HEAVY CHAIN DISEASE PROTEIN WIS IS SHOWN.  
 CC REF. 2.  
 CC -1- MISCELLANEOUS: THE SEQUENCE OF RESIDUES 42-76 WAS TAKEN FROM THE  
 CC REF. 2.  
 CC -1- MISCELLANEOUS: DISEASE PROTEIN WIS IS LACKING MOST OF THE V REGION  
 CC AND ALL OF THE CH1 REGION.  
 CC -1- MISCELLANEOUS: DISEASE PROTEIN ZUC LACK MOST OF THE V REGION, ALL  
 CC OF THE CH1 REGION, AND PART OF THE HINGE COMPARED WITH NORMAL  
 CC GAMMA-3 HEAVY CHAINS.  
 CC -1- MISCELLANEOUS: DISEASE PROTEIN OMM MAY REPRESENT AN ALLELIC FORM  
 CC OR ANOTHER GAMMA CHAIN SUBCLASS.  
 CC -1- MISCELLANEOUS: THE HINGE REGION IN GAMMA-3 CHAINS IS ABOUT FOUR  
 CC TIMES AS LONG AS IN OTHER GAMMA CHAINS AND CONTAINS THREE  
 CC IDENTICAL 15-RESIDUE SEGMENTS PRECEDED BY A SIMILAR 17-RESIDUE  
 CC SEGMENT (12-28).  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 CC EMBL; J00231; AAA52805.1; ALT\_SEQ.  
 DR PIR; A02149; G3HUWI.  
 DR HSP; P01857; IFC1.  
 DR Genew; HGNC:5527; IGHG3.  
 DR MIM; 147120; .  
 DR InterPro; IPR003006; Ig\_MHC.  
 DR InterPro; IPR003597; Ig-cl.  
 DR InterPro; IPR003600; Ig\_like.  
 DR Pfam; PF00047; ig; 2.  
 DR SMART; SM00410; Ig\_like; 1.  
 DR SMART; SM00407; IGH1; 1.  
 DR PROSITE; PS00290; IG\_MHC; 1.  
 KW Immunoglobulin domain; Immunoglobulin C region; Glycoprotein; Repeat.  
 FT DOMAIN 12 73  
 FT HINGE.  
 FT CH2.  
 FT CH3.  
 FT DOMAIN 74 183  
 FT DOMAIN 184 289  
 FT REPEAT 29 43  
 FT REPEAT 44 58  
 FT REPEAT 59 73  
 FT MOD\_RES 1 1  
 FT CARBOHYD 6 6  
 FT N-LINKED (GLCNAC...)  
 FT INTERCHAIN (WITH HEAVY CHAIN DIMER).  
 FT DISULFID 7 7  
 FT INTERCHAIN (WITH HEAVY CHAIN DIMER).  
 FT DISULFID 24 24  
 FT INTERCHAIN (WITH HEAVY CHAIN DIMER).  
 FT DISULFID 27 27  
 FT INTERCHAIN (WITH HEAVY CHAIN DIMER).  
 FT DISULFID 33 33

FT DISULFID 39 39 INTERCHAIN (WITH HEAVY CHAIN DIMER).  
 FT DISULFID 42 42 INTERCHAIN (WITH HEAVY CHAIN DIMER).  
 FT DISULFID 48 48 INTERCHAIN (WITH HEAVY CHAIN DIMER).  
 FT DISULFID 54 54 INTERCHAIN (WITH HEAVY CHAIN DIMER).  
 FT DISULFID 57 57 INTERCHAIN (WITH HEAVY CHAIN DIMER).  
 FT DISULFID 63 63 INTERCHAIN (WITH HEAVY CHAIN DIMER).  
 FT DISULFID 69 69 INTERCHAIN (WITH HEAVY CHAIN DIMER).  
 FT DISULFID 72 72 INTERCHAIN (WITH HEAVY CHAIN DIMER).  
 FT CARBOHYD 140 140 N-LINKED (GLCNAC...).  
 FT MOD\_RES 290 290 REMOVED POST-TRANSLATIONALLY.  
 FT VARIANT 126 127 QV -> EB (IN ZUC).  
 FT VARIANT 134 134 /FTID=VAR\_003890.  
 FT VARIANT 139 139 P -> L (IN OMM).  
 FT VARIANT 182 182 F -> Y (IN OMM).  
 FT VARIANT 182 182 /FTID=VAR\_003892.  
 FT VARIANT 227 227 T -> A (IN OMM).  
 FT VARIANT 227 227 /FTID=VAR\_003893.  
 FT VARIANT 227 227 S -> N (IN OMM).  
 FT VARIANT 227 227 /FTID=VAR\_003894.  
 FT VARIANT 279 279 MISSING (IN ZUC).  
 FT VARIANT 279 279 /FTID=VAR\_003895.  
 FT VARIANT 279 279 F -> Y (IN OMM).  
 FT VARIANT 279 279 /FTID=VAR\_003896.  
 SQ SEQUENCE 290 AA; 32331 MW; E69CB95705B2F46 CRC64;  
 Query Match 86.1%; Score 1148; DB 1; Length 290;  
 Best Local Similarity 74.8%; Pred. No. 2.1e-32;  
 Matches 211; Conservative 15; Mismatches 11; Indels 45; Gaps 1;  
 QY 10 VAELKAEAFMDKTH----- 24  
 : : : : :  
 Db 9 VSSELKTPGLDTHTCRCPEPKSCDTPPCRCPEPKSCDTPPCRCPEPKSCDTPPP 68  
 QY 25 CPPCPAPPELLGGPSVFLFPPPKKDTLMISRTPEVTVVVDVSHEDPEVKFMYVDGVVH 84  
 : : : : :  
 Db 69 CPRCPAPPELLGGPSVFLFPPPKKDTLMISRTPEVTVVVDVSHEDPEVKFMYVDGVVH 128  
 QY 85 NAKTKPREQVNSRYVSVLTVLHODWLNKGKVKVSNKALPAPIETKTSKAGOPRE 144  
 : : : : :  
 Db 129 NAKTKPREQVNSRYVSVLTVLHODWLNKGKVKVSNKALPAPIETKTSKAGOPRE 188  
 QY 145 PQVYTLPPSRDELTKNOVSLTCLVKGFYPSDIAVWESNGOPENNYKTPPVLDSDGSEFF 204  
 : : : : :  
 Db 189 PQVYTLPPSRDEMTKNOVSLTCLVKGFYPSDIAVWESSGOPENNYKTPPVLDSDGSEFF 248  
 QY 205 LYSKLTVDKSRWQGNVFCFSVMHEALHNHYTQKSLSLSPGK 246  
 : : : : :  
 Db 249 LYSKLTVDKSRWQGNVFCFSVMHEALHNHYTQKSLSLSPGK 290  
 RESULT 3  
 GC2\_HUMAN  
 ID GC2\_HUMAN STANDARD; PRT; 326 AA.  
 AC P01859;  
 DT 21-JUL-1986 (Rel. 01, Created)  
 DT 21-JUL-1986 (Rel. 01, Last sequence update)  
 DT 16-OCT-2001 (Rel. 40, Last annotation update)  
 DE Ig gamma-2 chain C region.  
 GN IGH2.  
 OS Homo sapiens (Human).  
 OC Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 OC NCBI\_TaxID=9606;  
 [1]  
 RN SEQUENCE OF 2-326 FROM N.A.  
 RX MEDLINE=82197621; PubMed=6804948;  
 RA Ellison J.W., Hood L.E.;  
 RT "Linkage and sequence homology of two human immunoglobulin gamma  
 RL heavy chain constant region genes.";  
 RL Proc. Natl. Acad. Sci. U.S.A. 79:1984-1988(1982).  
 [2]  
 RN SEQUENCE OF 88-115 FROM N.A.

RC TISSUE=Fetal liver;  
RX MEDLINE=83001943; PubMed=6811139;  
RA Takahashi N., Ueda S., Obata M., Nikaide T., Nakai S., Honjo T.;  
RT "Structure of human immunoglobulin gamma genes: implications for  
RT evolution of a gene family.";  
RL Cell 29:671-679(1982).  
RN [3]  
RP SEQUENCE OF 99-177 AND 310-326 FROM N.A.  
RX TISSUE=Fetal liver;  
RA MEDLINE=84235992; PubMed=6329676;  
RT Krawinkel U., Rabbits T.H.;  
RT "Comparison of the hinge-coding segments in human immunoglobulin gamma  
RT heavy chain genes and the linkage of the gamma 2 and gamma 4 subclass  
RT genes.";  
RL EMBO J. 1:403-407(1982).  
RN [4]  
RP SEQUENCE OF 1-325 (MYELOMA PROTEIN TIL).  
RX MEDLINE=81007873; PubMed=6774012;  
RA Wang A.-C., Tung E., Fudenberg H.H.;  
RT "The primary structure of a human IgG2 heavy chain: genetic,  
RT evolutionary, and functional implications.";  
RL J. Immunol. 125:1048-1054(1980).  
RN [5]  
RP SEQUENCE OF 1-85 AND 132-325 (MYELOMA PROTEIN ZIE).  
RX MEDLINE=80001357; PubMed=113060;  
RA Connell G.E., Parr D.M., Hofmann T.;  
RT "The amino acid sequences of the three heavy chain constant region  
RT domains of a human IgG2 myeloma protein.";  
RL Can. J. Biochem. 57:758-767(1979).  
RN [6]  
RP SEQUENCE OF 238-275 (ZIE).  
RX MEDLINE=80114419; PubMed=118920;  
RA Hofmann T., Parr D.M.;  
RT "A note of the amino acid sequence of residues 381-391 of human  
RT immunoglobulin gamma chains.";  
RL Mol. Immunol. 16:923-925(1979).  
RN [7]  
RP REVISIONS TO 25; 59; 60 AND 264-268 (ZIE).  
RA Hofmann T., Parr D.M.;  
RT Submitted (MAR-1980) to the PIR data bank.  
RN [8]  
RP SEQUENCE OF 1-121 (DOT).  
RX MEDLINE=95255298; PubMed=7737190;  
RA Stoppini M., Bellotti V., Negri A., Merlini G., Garver F., Ferri G.;  
RT "Characterization of the two unique human anti-flavin monoclonal  
RT immunoglobulins.";  
RL Eur. J. Biochem. 228:885-893(1995).  
RN [9]  
RP DISULFIDE BONDS.  
RX MEDLINE=72033500; PubMed=4940472;  
RA Milstein C., Frangione B.;  
RT "Disulphide bridges of the heavy chain of human immunoglobulin G2.";  
RL Biochem. J. 121:217-225(1971).  
RN [10]  
RP DISULFIDE BONDS.  
RX MEDLINE=69064124; PubMed=5782707;  
RA Frangione B., Milstein C., Pink J.R.L.;  
RT "Structural studies of immunoglobulin G.";  
RL Nature 221:145-148(1969).  
RN [11]  
RP This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
RN [12]  
DR EMBL; J00230; AAB59393.1; -;  
DR PIR; A02148; G2HU.  
DR HSSP; P01857; 1FC1.  
DR Genew; HGNC:5526; IGHG2.  
DR MM; 147110; -;

DR InterPro; IPR003006; Ig\_MHC.  
DR InterPro; IPR003597; Ig\_c1.  
DR InterPro; IPR003600; Ig\_like.  
DR Pfam; PF00047; Ig; 3.  
DR SMART; SM00410; Ig\_like; 1.  
DR SMART; SM00407; Igc1; 2.  
DR PROSITE; PS00290; IG\_MHC; 2.  
KW Immunoglobulin domain; Immunoglobulin C region.  
FT DOMAIN 1 98 CH1.  
FT DOMAIN 99 110 HINGE.  
FT DOMAIN 111 219 CH2.  
FT DOMAIN 220 326 CH3.  
FT DISULFID 14 14  
FT DISULFID 27 83  
FT DISULFID 102 102  
FT DISULFID 103 103  
FT DISULFID 106 106  
FT DISULFID 109 109  
FT DISULFID 140 200  
FT DISULFID 246 304  
FT SITE 156 156  
FT MOD\_RES 326 326  
FT VARIANT 60 60  
FT CONFLICT 109 109  
SQ SEQUENCE 326 AA; 35884 MW; 8310878C6878C9C CRC64;  
Query Match 86.1%; Score 1148; DB 1; Length 326;  
Best Local Similarity 84.1%; Pred. No. 3e-32; Indels 24; Gaps 2;  
Matches 212; Conservative 10; Mismatches 6;  
QY 18 FMDKTHT-----CPPCPAPPELLGGPSVFLFPPKPKDTLMISR 54  
DB 76 FGQTYTCNVNDHKPSNTKVKDKTVERKCCVECPCCAPP-VAGPSVFLFPPKPKDTLMISR 134  
QY 55 TPEVTCVVVDVSHEDPEVKFNQYVDGVEVHNNAKTPRREEQYNSTRVSVLVFLVHODWLN 114  
DB 135 TPEVTCVVVDVSHEDPEVQFNQYVDGVEVHNNAKTPRREEQFNSTRVSVLVFLVHODWLN 194  
QY 115 GKEYCKVSNKALPAPIEKTISKAKGQPREPOVYTLPPSRDELTKNOVSLTCLVKGFYPS 174  
DB 195 GKEYCKVSNKALPAPIEKTISKAKGQPREPOVYTLPPSRDELTKNOVSLTCLVKGFYPS 254  
QY 175 DIAVWESNGQPENNYKTTPPVLDSDGSFPLYSKLTVDKSRWQQGQNVFSCSYNHEALHNN 234  
DB 255 DIAVWESNGQPENNYKTTPPVLDSDGSFPLYSKLTVDKSRWQQGQNVFSCSYNHEALHNN 314  
QY 235 YTKQSLSLSPGK 246  
DB 315 YTKQSLSLSPGK 326  
RESULT 4  
GC4\_HUMAN  
ID GC4\_HUMAN STANDARD; PRT; 327 AA.  
AC P01861;  
DT 21-JUL-1986 (Rel. 01, Created)  
DT 21-JUL-1986 (Rel. 01, Last sequence update)  
DT 16-OCT-2001 (Rel. 40, Last annotation update)  
DE Ig gamma-4 chain C region.  
GN IGHG4.  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
OX NCBI\_TaxID=9606;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=83157104; PubMed=6299662;  
RA Ellison J.W., Buxbaum J.N., Hood L.E.;  
RT "Nucleotide sequence of a human immunoglobulin C gamma 4 gene.";  
RL DNA 1:11-18(1981).  
RN [2]

```
RP SEQUENCE OF 1-30 AND 81-326.
RX MEDLINE=70207560; PubMed=4192699;
RA Pink J.R.L., Buttery S.H., de Vries G.M., Millstein C.;
RT "Human immunoglobulin subclasses. Partial amino acid sequence of the
RL Biochem. J. 117:33-47(1970).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; K01316; AAB59394.1; ALT_INIT.
DR PIR; A02150; G4HU.
DR HSSP; P01842; 7FAB.
DR Genew; HGNC:5528; IGHG4.
DR MIM; 147130; -.
DR InterPro; IPR003006; Ig_MHC.
DR InterPro; IPR003597; Ig_C1.
DR InterPro; IPR003600; Ig_Like.
DR Pfam; PF00047; Ig; 3.
DR SMART; SM00410; Ig_Like; 1.
DR SMART; SM00407; Igc1; 2.
DR PROSITE; PS00290; IG_MHC; 2.
KW Immunoglobulin domain; Immunoglobulin C region.
FT DOMAIN 1 1
FT 1 CH1.
FT 98 HINGE.
FT 110 HINGE.
FT 111 CH2.
FT 220 CH2.
FT 221 CH3.
FT 327 CH3.
FT 328 INTERCHAIN (WITH A LIGHT CHAIN).
FT 329 INTERCHAIN (WITH A HEAVY CHAIN).
FT 330 INTERCHAIN (WITH A HEAVY CHAIN).
FT 331 DISULFID 14 14
FT 332 DISULFID 27 83
FT 333 DISULFID 106 106
FT 334 DISULFID 109 109
FT 335 DISULFID 141 201
FT 336 DISULFID 247 305
SQ SEQUENCE 327 AA; 35940 MW; 3EBDB811EF208E7A CRC64;
Query Match 85.8%; Score 1144; DB 1; Length 327;
Best Local Similarity 84.7%; Pred. No. 4.1e-32;
Matches 211; Conservative 9; Mismatches 6; Indels 23; Gaps 1;
QY 21 KTHT-----CPCPAPELLGGPSVFLFPPKPKDTLMISRTPE 57
Db 79 KTYCNVDHKSPTKVKRVESKYGPPCPAPEFLGGPSVFLFPPKPKDTLMISRTPE 138
QY 58 VTCVVYDVSHEDPEVAFNNYVDGVEYVHNAKTPREQYNSTYRVSVLTVLHODWLNKGE 117
Db 139 VTCVVYDVSDQDEPEVQFNYYVDGVEYVHNAKTPREQFNSTYRVSVLTVLHODWLNKGE 198
QY 118 YKCKVSNKALPAPIETKISKAKQPREPVYTLPPSDELTKNQVSLTCLVKGFYPSDIA 177
Db 199 YKCKVSNKGLPSSIEKISKAKQPREPVYTLPPSDEETKQNQVSLTCLVKGFYPSDIA 258
QY 178 VEVESNGQPNKYKTPPVLDSDGSFELYSKLTVDKSRWQGNVFCSVNHEALHNYHQ 237
Db 259 VEVESNGQPNKYKTPPVLDSDGSFELYSLRVTVDKSRWQGNVFCSVNHEALHNYHQ 318
QY 238 KSLSLSPGK 246
Db 319 KSLSLSLGK 327
RESULT 5
GC_RABIT STANDARD; PRT; 323 AA.
AC P01870;
DT 21-JUL-1986 (Rel. 01, Created)
DT 21-JUL-1986 (Rel. 01, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
```

```
DE Ig gamma chain C region.
OS Eryctolagus cuniculus (Rabbit).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Lagomorpha; Leporidae; Oryctolagus.
OX NCBI_TaxID=9986;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=84030930; PubMed=6313520;
RA Bernstein K.E., Alexander C.B., Mage R.G.;
RT "Nucleotide sequence of a rabbit IgG heavy chain from the recombinant
RT F-1 haplotype.";
RL Immunogenetics 18:387-397(1983).
RN [2]
RP SEQUENCE OF 1-128.
RX MEDLINE=76135469; PubMed=1243651;
RA Pratt D.M., Mole L.E.;
RT "Sequence studies on the constant region of the Fd sections of rabbit
RT immunoglobulin G of different allotype.";
RL Biochem. J. 151:337-349(1975).
RN [3]
RP SEQUENCE OF 88-266 FROM N.A.
RX MEDLINE=83299917; PubMed=6193512;
RA Martens C.L., Moore K.W., Steinmetz M., Hood L., Knight K.L.;
RT "Heavy chain genes of rabbit IgG: isolation of a cDNA encoding gamma
RT heavy chain and identification of two genomic C gamma genes.";
RL Proc. Natl. Acad. Sci. U.S.A. 79:6018-6022(1982).
RN [4]
RP SEQUENCE OF 132-161.
RX MEDLINE=70110015; PubMed=5461106;
RA Fruchter R.G., Jackson S.A., Mole L.E., Porter R.R.;
RT "Sequence studies of the Fd section of the heavy chain of rabbit
RT immunoglobulin G.";
RL Biochem. J. 116:249-259(1970).
RN [5]
RP SEQUENCE OF 129-131 AND 155-322.
RA Hill R.L., Lebovitz H.E., Fellows R.E. Jr., Delaney R.;
RL (In) Killander J. (eds.);
RL Gamma globulins, Nobel symp. 3, pp.109-127, Almqvist and Wiksell,
RL Stockholm (1967).
CC -!- MISCELLANEOUS: REF.1 SEQUENCE HAS THE D12 ALLOTYPE MARKER,
CC 104-THR, AND THE E14 MARKER, 185-THR. REF.3 HAS THE D11 AND E15
CC MARKERS AND REF.5 THE E15 MARKER.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; M16426; AAA31289.1; -.
DR PIR; A02161; GHRB.
DR HSSP; P01857; 1FC1.
DR InterPro; IPR003006; Ig_MHC.
DR InterPro; IPR003597; Ig_C1.
DR Pfam; PF00047; Ig; 2.
DR SMART; SM00407; Igc1; 2.
DR PROSITE; PS00290; IG_MHC; 1.
KW Immunoglobulin domain; Immunoglobulin C region.
FT NON_TER 1 1
FT 1 VARIANT 104 104
FT 185 T -> M (IN D11 MARKER).
FT 185 T -> A (IN E15 MARKER).
FT 48 N -> E (IN REF. 2).
FT 71 V -> VPV (IN REF. 2).
FT 144 Q -> E (IN REF. 3 AND 4).
FT 173 N -> D (IN REF. 5).
FT 187 Q -> E (IN REF. 3 AND 5).
FT 201 N -> D (IN REF. 5).
FT 218 N -> Q (IN REF. 5).
FT 233 E -> Q (IN REF. 5).
FT 246 N -> D (IN REF. 5).
FT 256 E -> G (IN REF. 5).
```

```

FT CONFLICT 260 260 N -> D (IN REF. 5).
FT CONFLICT 266 266 N -> D (IN REF. 5).
FT CONFLICT 280 280 Y -> W (IN REF. 5).
FT CONFLICT 284 284 N -> S (IN REF. 5).
SQ SEQUENCE 323 AA; 35404 MW; 69E8AAL18D579A8B CRC64;

Query Match 69.2%; Score 923; DB 1; Length 323;
Best Local Similarity 69.8%; Pred. No. 1.7e-24;
Matches 169; Conservative 29; Mismatches 39; Indels 5; Gaps 2;

QY 10 VAEKLKEAFMDKT---HTC--PPCAPPELLGGPSVFLPPKPKDITLMISRTPEVTCVVVD 64
DB 82 VAHPATNTKVDKTVAPSTCKTPCPPELLGGPSVFLPPKPKDITLMISRTPEVTCVVVD 141
QY 65 VSHEDPEVKFNMYVDGVEVHNKTPREQYNSTYRVSVLVHLQDNLNGREYKCKVSN 124
DB 142 VSQDDPEVQFTWYINNEQYRTARPLREQQFNSTIRVSVTLTIHQDNLNGREYKCKVHN 201
QY 125 KALPAPIETISKAGQPPREPOVYTLPPSRDELTKNOVSLTCLVKGFPSPDIKAVESWG 184
DB 202 KALPAPIETISKAGQPPLEKPYVTMGPPREELSSRSVSLTCMNGFYPSDISVEWKG 261
QY 185 QPENNYKTPPVLDSDGSPFLYSKLTVDKSRWQGNVFCSCVMHEALHNHYTKSLSLSP 244
DB 262 KAEDNKTTPPAVLDSGSLYLNKUSVPTSEWQGDVFTCSVMHEALHNHYTKSISRSP 321
QY 245 GK 246
DB 322 GK 323

RESULT 6
GC2_CAVPO STANDARD; PRT; 329 AA.
AC P01862;
DT 21-JUL-1986 (Rel. 01, Created)
DT 15-JUL-1999 (Rel. 38, Last sequence update)
DE Ig gamma-2 chain C region.
OS Cavia porcellus (Guinea pig).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Hystriognathi; Caviidae; Cavia.
OX NCBI_TaxID=10141;
RN [1]
RP SEQUENCE OF 1-3.
RA Trischmann T.M.;
RL Submitted (APR-1975) to the PIR data bank.
RN [2]
RP SEQUENCE OF 4-68.
RX MEDLINE=71058471; PubMed=5538606;
RA Birshstein B.K.; Hussain Q.Z.; Cebra J.J.;
RT "Structure of heavy chain from strain 13 guinea pig
RT immunoglobulin-G(2). 3. Amino acid sequence of the region around the
RT half-cystine joining heavy and light chains.";
RL Biochemistry 10:18-25(1971).
RN [3]
RP SEQUENCE OF 69-133 AND 312-329.
RX MEDLINE=71058486; PubMed=5538616;
RA Turner K.J.; Cebra J.J.;
RT "Structure of heavy chain from strain 13 guinea pig
RT immunoglobulin-G(2). II. Amino acid sequence of the carboxyl-terminal
RT and hinge region cyanogen bromide fragments.";
RL Biochemistry 10:9-17(1971).
RN [4]
RP SEQUENCE OF 134-226.
RX MEDLINE=75036072; PubMed=4429665;
RA Tracey D.E.; Cebra J.J.;
RT "Primary structure of the CH2 homology region from guinea pig IgG2
RT antibodies.";
RL Biochemistry 13:4796-4803(1974).
RN [5]
RP SEQUENCE OF 227-311.
RX MEDLINE=75036073; PubMed=4609467;

```

```

RA Trischmann T.M.; Cebra J.J.;
RT "Primary structure of the CH3 homology region from guinea pig IgG2
RT antibodies.";
RL Biochemistry 13:4804-4811(1974).
RN [6]
RP DISULFIDE BONDS.
RX MEDLINE=71058474; PubMed=4922544;
RA Oliveira B.; Lamm M.E.;
RT "Interchain disulfide bridges of guinea pig gamma-2-immunoglobulin.";
RL Biochemistry 10:26-31(1971).
CC -1- MISCELLANEOUS: THIS CHAIN WAS ISOLATED FROM POOLED SERUM OF STRAIN
CC 13 INBRED GUINEA PIGS.
DR PIR; A02151; G2GP.
DR HSSP; P01842; 7FAB.
DR InterPro; IPR003006; Ig_MHC.
DR InterPro; IPR003597; Ig_c1.
DR InterPro; IPR003600; Ig_Like.
DR Pfam; PF00047; Ig; 2.
DR SMART; SM00410; Ig_Like; 1.
DR SMART; SM00407; IGC1; 2.
DR PROSITE; PS00290; IG_MHC; 1.
KW Immunoglobulin domain; Immunoglobulin C region; Glycoprotein.
FT NOX_TER 1 1
FT DISULFID 16 16 INTERCHAIN (WITH A LIGHT CHAIN).
FT DISULFID 28 79
FT DISULFID 105 105 INTERCHAIN (WITH A HEAVY CHAIN).
FT DISULFID 107 107 INTERCHAIN (WITH A HEAVY CHAIN).
FT DISULFID 110 110 INTERCHAIN (WITH A HEAVY CHAIN).
FT DISULFID 142 202
FT CARBOHYD 178 178 N-LINKED (GLCNAC. . .).
FT DISULFID 248 308
SQ SEQUENCE 329 AA; 36074 MW; 5D231B7164D1FBA9 CRC64;

Query Match 67.4%; Score 899; DB 1; Length 329;
Best Local Similarity 67.1%; Pred. No. 1.2e-23;
Matches 167; Conservative 25; Mismatches 44; Indels 13; Gaps 2;

QY 10 VAEKLKEAFMDKT---HTC--PPCAPPELLGGPSVFLPPKPKDITLMISRTPEV 58
DB 81 VAHPASSYTKVDKTVPEIRTPZPBCTCPKCPPELGGPSVFLPPKPKDITLMISLTPRV 140
QY 59 TCVVVDVSHEDPEVKFNMYVDGVEVHNKTPREQYNSTYRVSVLVHLQDNLNGREY 118
DB 141 TCVVVDVSDQDEPEVQFTWFDNKPVGNAETKPRVQYNTTFRVESVLPITQHDNLGRKEF 200
QY 119 KCKYSNKALPAPIETISKAGQPPREPOVYTLPPSRDELTKNOVSLTCLVKGFPSPDIKAV 178
DB 201 KCKVYNKALPAPIETISKAGQPPREPOVYTLPPSRDELTKNOVSLTCLVKGFPSPDIKAV 260
QY 179 EWESNGQP--ENNYYKTPPVLDSDGSPFLYSKLTVDKSRWQGNVFCSCVMHEALHNHYT 236
DB 261 EWASNRVPVSEKEYKNTPTIEDADGSLYSLYSLYSLYSLYSLYSLYSLYSLYSLYSLY 320
QY 237 QKSLSLSPG 245
DB 321 QKAISRSPG 329

RESULT 7
GC3_MOUSE STANDARD; PRT; 329 AA.
AC P22436;
DT 01-AUG-1991 (Rel. 19, Created)
DT 01-AUG-1991 (Rel. 19, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Ig gamma-3 chain C region, secreted form.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=85027161; PubMed=6092053;

```

```

RA Wels J.A., Word C.J., Rimm D., Der-Balan G.P., Martinez H.M.,
RA Tucker P.W., Blattner F.R.;
RT "Structural analysis of the murine IgG3 constant region gene.";
RL EMBO J. 3:2041-2046(1984).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: J00451; ; NOT_ANNOTATED_CDS.
DR PIR: B02156; G3MSC.
DR HSP: P01857; JFCL.
DR InterPro: IPR003006; Ig_MHC.
DR InterPro: IPR003597; Ig-cl.
DR Pfam: PF000360; Ig_like.
DR SMART: SM00407; Ig; 3.
DR SMART: SM00410; Ig_like; 1.
DR PROSITE: PS00290; IG_MHC; 1.
KW Immunoglobulin domain; Immunoglobulin C region.
FT NON_TER 1 97 CH1.
FT DOMAIN 98 113 HINGE.
FT DOMAIN 114 223 CH2.
FT DOMAIN 224 327 CH3.
SQ SEQUENCE 329 AA; 36228 MW; F45827174182BAD6 CRC64;

Query Match 63.9%; Score 853; DB 1; Length 329;
Best Local Similarity 55.8%; Pred. No. 4.7e-22;
Matches 155; Conservative 37; Mismatches 45; Indels 41; Gaps 3;

QY 10 VAEKLEAFM-----DKTHTC-----PP--C 28
Db 52 VSSVLQSGFYSLSLVTPSPSTWPSQTVICVAHPASKTELKRIEPRKPSPPGSSC 111
QY 29 PAPELIGGPSVFLPPPKDTLMISRTPEVTCVVVDVSHEDPEVKFNNYVGVGEVHNAKT 88
Db 112 PPGNIGGPSVFIIPPKPKDLMISLTPKVTVCVVVDVSEDDPDVHVSFVDNKEVHTAWT 171
QY 89 KPREQYNSTYRVVSVLTVLHQDWLNGKEYKCKVSNKALPAPIEKTTISKAKGQPREPQVY 148
Db 172 QPREAQVNSYFVVSVLPIQHODMRCGEKCKVNNKALPAPIERTISKPKGRAQTQVY 231
QY 149 TLPPSRDELTKNOVSLTCLVKGYFYPDSIAVWESNGOPENNYKTTTPVLDSDGSFFLYSK 208
Db 232 TIPPPEQMSKKVSLTCLVTNPFSEISVSEWERNGELEQDYKNTPTPLDSDGYFLYLYSK 291
QY 209 LTVDKSRWQGNVFCSCVMHEALHNHYTQKSLSLSPCK 246
Db 292 LTVDTDSWLOGEIFTCSVWHEALHNHYTQKSLSRSPCK 329

RESULT 8
GCB_RAT
ID GCB_RAT STANDARD; PRT; 333 AA.
AC P20761;
DT 01-FEB-1991 (Rel. 17, Created)
DT 01-FEB-1991 (Rel. 17, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE Ig gamma-2B chain C region.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=89232738; PubMed=3149946;
RA Brueggemann M.;

"Evolution of the rat immunoglobulin gamma heavy-chain gene family.";
RL Gene 74:473-482(1988).
RT PIR: PS0018; PS0018.
DR HSP: P01842; 7FAB.
DR InterPro: IPR003006; Ig_MHC.
DR InterPro: IPR003597; Ig-cl.
DR Pfam: PF000360; Ig_like.
DR SMART: SM00407; Ig; 3.
DR SMART: SM00410; Ig_like; 1.
DR PROSITE: PS00290; IG_MHC; 1.
KW Immunoglobulin domain; Immunoglobulin C region.
FT NON_TER 1 15 INTERCHAIN (WITH A LIGHT CHAIN).
FT DISULFID 15 15
FT DISULFID 27 80
FT DISULFID 106 106 INTERCHAIN (WITH A HEAVY CHAIN).
FT DISULFID 109 109 INTERCHAIN (WITH A HEAVY CHAIN).
FT DISULFID 112 112 INTERCHAIN (WITH A HEAVY CHAIN).
FT DISULFID 115 115 INTERCHAIN (WITH A HEAVY CHAIN).
FT DISULFID 147 207
FT DISULFID 253 311
SQ SEQUENCE 333 AA; 36497 MW; 55F8B64D48D460A6 CRC64;

Query Match 63.6%; Score 849; DB 1; Length 333;
Best Local Similarity 63.2%; Pred. No. 6.8e-22;
Matches 153; Conservative 34; Mismatches 52; Indels 3; Gaps 1;

QY 8 DKVAELKEAFMDKTHTCP---PCPAPELIGGPSVFLPPPKDTLMISRTPEVTCVVVD 64
Db 92 DKYVRRNGGIGHKPTCTCHKCPVPELLGGPSVFIIPPKPKDILLISQNAKVTGVVD 151
QY 65 VSHEDPEVKFNWYVDGVEVHNAKTKPREEQYNSTYRVVSVLTVLHQDWLNGKEYKCKVSN 124
Db 152 VSEEDPDVQSFVNNVEVHTAQTPREEQYNSTFRVVSALPIQHODMRCGEKCKVNN 211
QY 125 KALPAPIEKTISKAKGQPREPQVYTLPPSRDELTKNOVSLTCLVKGYFYPDSIAVWESNG 184
Db 212 KALPSPTEKTISKPKGLVRKPVVYMGPPTEBOLTEQVSLTCLTSGFLPDIGVETNSG 271
QY 185 QPENNYKTTTPVLDSDGSFFLYSKLTVDKSRWQGNVFCSCVMHEALHNHYTQKSLSLSP 244
Db 272 HIEKNYKNTPEVMDSDGSFFMYSKLNVRSRWDSPAPFCVSVVHEGLHNHVSISRP 331
QY 245 GK 246
Db 332 GK 333

RESULT 9
GCB_MOUSE
ID GCB_MOUSE STANDARD; PRT; 398 AA.
AC P03987;
DT 23-OCT-1986 (Rel. 02, Created)
DT 01-AUG-1991 (Rel. 19, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE Ig gamma-3 chain C region, membrane-bound form.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=85027161; PubMed=6092053;
RA Wels J.A., Word C.J., Rimm D., Der-Balan G.P., Martinez H.M.,
RA Tucker P.W., Blattner F.R.;
RT "Structural analysis of the murine IgG3 constant region gene.";
RL EMBO J. 3:2041-2046(1984).
[2]
RP SEQUENCE OF 328-398 FROM N.A.
RX MEDLINE=8401483; PubMed=6314258;
RA Komaromy M., Clayton L., Rogers J., Robertson S., Kettman J.,
RA Wall R.;
RT "The structure of the mouse immunoglobulin in gamma 3 membrane gene

```

```

RT segment.";
RL Nucleic Acids Res. 11:6775-6785(1983).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; J00451; AAB59655.1; -
DR EMBL; V01526; CAA24767.1; ALT_SEQ.
DR PIR; A02155; G3MSM.
DR HSP; P01857; 1FC1.
DR InterPro; IPR003006; Ig_MHC.
DR InterPro; IPR003597; Ig_c1.
DR Pfam; PF00047; Ig; 3.
DR SMART; SM00410; Ig-like; 1.
DR SMART; SM00407; IGc1; 2.
DR PROSITE; PS00290; IG_MHC; 1.
KW Immunoglobulin domain; Immunoglobulin C region; Glycoprotein;
KW Transmembrane; Alternative splicing.
FT NON_TER 1
FT DOMAIN 1 97 CH1.
FT DOMAIN 98 113 HINGE.
FT DOMAIN 114 223 CH2.
FT DOMAIN 224 327 CH3.
FT TRANSMEM 346 362 POTENTIAL.
FT DOMAIN 363 398 CYTOPLASMIC (POTENTIAL).
FT CONFLICT 333 333 E -> G (IN REF. 2).
FT CONFLICT 342 342 E -> O (IN REF. 2).
FT CONFLICT 388 388 P -> F (IN REF. 2).
SQ SEQUENCE 398 AA; 43929 MW; C7F264B50A41B95 CRC64;

Query Match 63.1%; Score 842; DB 1; Length 398;
Best Local Similarity 51.3%; Pred. NO. 2e-21;
Matches 154; Conservative 41; Mismatches 47; Indels 58; Gaps 4;

QY 3 LKAFYDK-----VAEKLKEAFM-----DKTHFC----- 25
DB 28 VAGIEPEPVTKWNGALSSGVRTVSSVLQSGFYSLSLVTVPSSVTSQVICNVAHPA 87
QY 26 -----PP-----CPAPELLGSPVFLPPPKPKDTLMISRTPEVTCVVVD 64
DB 88 SKTELKRIEPRIPKSTPPGSSCPPGNILGSPVFIIPPCKDALMSLPKVTVCVVVD 147
QY 65 VSHEDPEVKFNVDGVEVHNKTKPREEQNSTYRVVSVLTVLHQDWLNGKEYCKKYSN 124
DB 148 VSEDDPDVHVSFWFDNKEVHTAQTQPREAQYNSTPRVYSALPIQHQQDMRGKFEKCKVNN 207
QY 125 KALPAPIETKISKAKQPREPOVYTLPPSRDELTKNOVSLTCLVKGFYPSDIAVWESNG 184
DB 208 KALPAPIETKISKAGRAQTQPVYITPPRQMSKRVSLTCLVTFNFFSEALISVEWERNG 267
QY 185 QPENNYKTPPVLDSGDFLYSKLTVDKSRMQQGNFSCSVMHAEALHNHYTQKSLSP 244
DB 268 ELEQDYKNTPTPLDSDGTGYFLYSKLTVDTSNLOGEIFTCVSVHAEALHNHHTQKNLSRP 327

RESULT 10
GCL_RAT
ID GCL_RAT STANDARD; PRT; 326 AA.
AC P20759;
DT 01-FEB-1991 (Rel. 17, Created)
DT 01-FEB-1991 (Rel. 17, Last sequence update)
DT 15-JUL-1999 (Rel. 36, Last annotation update)
DE Ig gamma-1 chain C region.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Rattus.
OX NCBI_TaxID=10116;

```

```

RN SEQUENCE FROM N.A.
RP MEDLINE=89232738; PubMed=3149946;
RA Brueggemann M.;
RT "Evolution of the rat immunoglobulin gamma heavy-chain gene family.";
RL Gene 74:473-482(1988).
DR PIR; PS0017; PS0017.
DR HSP; P01842; 7FAB.
DR InterPro; IPR003006; Ig_MHC.
DR InterPro; IPR003597; Ig_c1.
DR Pfam; PF00047; Ig; 3.
DR SMART; SM00407; IGc1; 2.
DR PROSITE; PS00290; IG_MHC; 1.
KW Immunoglobulin domain; Immunoglobulin C region; Glycoprotein.
FT NON_TER 1
FT DOMAIN 1 97 CH1.
FT DOMAIN 98 112 HINGE.
FT DOMAIN 113 219 CH2.
FT DOMAIN 220 326 CH3.
FT DISULFID 27 82
FT DISULFID 102 102 INTERCHAIN (WITH A HEAVY CHAIN).
FT DISULFID 106 106 INTERCHAIN (WITH A HEAVY CHAIN).
FT DISULFID 109 109 INTERCHAIN (WITH A HEAVY CHAIN).
FT DISULFID 111 111 INTERCHAIN (WITH A HEAVY CHAIN).
FT DISULFID 140 200
FT DISULFID 246 304
FT CARBOHYD 176 176 N-LINKED (GLCNAC...) (POTENTIAL).
SQ SEQUENCE 326 AA; 35946 MW; 013BAB45EF49B9DA CRC64;

Query Match 61.9%; Score 826; DB 1; Length 326;
Best Local Similarity 59.1%; Pred. NO. 4e-21;
Matches 146; Conservative 42; Mismatches 45; Indels 14; Gaps 3;

QY 10 VAEKLKEAFMDKT-----HTCPCPAPELLGG---PSVLEFPKPKDTLMISTPEVT 59
DB 84 VAHPASSTKVDKIVPRNCGGCKPC-----ICTSEVSSVFIPPKPKDVLITITPKVT 139
QY 60 CVVYDVSHEDPEVKFNVDGVEVHNKTKPREEQNSTYRVVSVLTVLHQDWLNGKEYK 119
DB 140 CVVVDISODDPEVHFSEWFDVVEVHTAQTTPPEEQNFSTERSVSELPILHQDWLNGRTFR 199
QY 120 CKVSNKALPAPIETKISKAKQPREPOVYTLPPSRDELTKNOVSLTCLVKGFYPSDIAVE 179
DB 200 CKVTSAAAPSPSIEKTSKPEGRQVPHVYITMSPTKEEMTQNEVITCVKGFYPPDIYE 259
QY 180 WESNGQPNYKTPPVLDSGDFLYSKLTVDKSRMQQGNFSCSVMHAEALHNHYTQKS 239
DB 260 WQMGQPOENYKNTPTTMDTDCSYFLYSKLVKVKKEKQOQGNFTTCSVLHEGLHNHHTKS 319
QY 240 LSLSPGK 246
DB 320 LSHSPGK 326

RESULT 11
GCL_MOUSE
ID GCL_MOUSE STANDARD; PRT; 324 AA.
AC P01868;
DT 21-JUL-1986 (Rel. 01, Created)
DT 21-JUL-1986 (Rel. 01, Last sequence update)
DT 30-MAY-2000 (Rel. 39, Last annotation update)
DE Ig gamma-1 chain C region.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN SEQUENCE FROM N.A.
RP MEDLINE=80045036; PubMed=115593;
RA Honjo T., Obata M., Yamawaki-Kataoka Y., Kataoka T., Kawakami T.,
RA Takahashi N., Mano Y.;
RT "Cloning and complete nucleotide sequence of mouse immunoglobulin
gamma 1 chain gene.";

```



```

Cell 18:559-568(1979).
[2]
SEQUENCE OF 76-324 FROM N.A. (MYELOMA PROTEIN MOPC 31C).
RX MEDLINE=80202559; PubMed=6769752;
RA Obata M., Tamawaki-Kataoka Y., Takahashi N., Kataoka T., Shimizu A.,
RA Mano Y., Selman J.G., Peterlin B.M., Leder P., Honjo T.;
RT "Immunoglobulin gamma 1 heavy chain gene: structural gene sequences
RT cloned in a bacterial plasmid.";
RL Gene 9:87-97(1980).
RN [3]
RN SEQUENCE OF 70-322 FROM N.A. (MYELOMA PROTEIN MOPC 21).
RX MEDLINE=80012837; PubMed=113776;
RA Rogers J., Clarke P., Salser W.;
RA "Sequence analysis of cloned cDNA encoding part of an immunoglobulin
RT heavy chain.";
RL Nucleic Acids Res. 6:3305-3321(1979).
RN [4]
RN SEQUENCE (MYELOMA PROTEIN MOPC 21).
RX MEDLINE=78242288; PubMed=98524;
RA Adetugbo K.;
RA "Evolution of immunoglobulin subclasses. Primary structure of a
RT murine myeloma gammal chain.";
RL J. Biol. Chem. 253:6068-6075(1978).
RN [5]
RN DISULFIDE BONDS (MOPC 21).
RX MEDLINE=73008889; PubMed=5073237;
RA Svasti J., Milstein C.;
RT "The disulphide bridges of a mouse immunoglobulin G1 protein.";
RL Biochem. J. 126:837-850(1972).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; V00793; CAA24172.1; -
DR EMBL; V00793; CAA24173.1; -
DR EMBL; V00793; CAA24174.1; -
DR EMBL; V00793; CAA24175.1; -
DR EMBL; V00795; CAA24176.1; -
DR PIR; A02159; GLMS.
DR HSSP; P01842; 7FAB.
DR GlycoSuiteDB; P01868; -
DR MGD; MGI:96446; Igh-4.
DR InterPro; IPR003006; Ig_MHC.
DR InterPro; IPR003597; Ig_cl.
DR Pfam; PF00047; Ig; 3.
DR SMART; SM00407; Igcl; 2.
DR PROSITE; PS00290; IG_MHC; 1.
KW Immunoglobulin domain; Immunoglobulin C region; Glycoprotein;
KW Alternative splicing.
FT NON_TER 1 97
FT DOMAIN 1 97 CHI.
FT DOMAIN 98 110 HINGE.
FT DOMAIN 111 217 HINGE.
FT DOMAIN 218 324 CH2.
FT DISULFID 27 82 CH3.
FT DISULFID 102 102 INTERCHAIN (WITH A LIGHT CHAIN).
FT DISULFID 104 104 INTERCHAIN (WITH A HEAVY CHAIN).
FT DISULFID 107 107 INTERCHAIN (WITH A HEAVY CHAIN).
FT DISULFID 109 109 INTERCHAIN (WITH A HEAVY CHAIN).
FT DISULFID 138 198
FT CARBOHYD 174 174 N-LINKED (GLCNAC. . .).
FT FTID=CAR_000055.
FT DISULFID 244 302
FT MOD_RES 324 324 REMOVED POST-TRANSLATIONALLY.
FT CONFLICT 276 276 N -> D (IN REF. 3).
FT CONFLICT 278 278 N -> D (IN REF. 3).
SQ SEQUENCE 324 AA; 35704 MW; A338812F3D1F2C93 CRC64;

```

```

Query Match 61.7%; Score 823; DB 1; Length 324;
Best Local Similarity 59.7%; Pred. No. 4.9e-21;
Matches 141; Conservative 48; Mismatches 42; Indels 5; Gaps 2;

QY 13 KLEAFPMKTHHTCPP--CPAPELLGGPSVFLFPKPKDITLMISRTPEVTCVVVDVSHEDP 70
   : : : | | | : : : : : : : : : : : : : : : : : : : : : : : :
DB 92 KVDKTIIVPRDCGCKPCICTVPEV---SSVFIFPPKPKDVLITITPKVTCVVVDISKDDP 148
   : : : | | | : : : : : : : : : : : : : : : : : : : : : : : :
QY 71 EVKFNMYVDGVEVHNAKTKPREQYNSTYRVSVLTVTLHODWNLGKEYCKVSNKALPAP 130
   : : : | | | : : : : : : : : : : : : : : : : : : : : : : : :
DB 149 EVQFSWFDVDDVEVHTAQTQPREEQNFTRSVSELPIMHODWNLGKFKCRVNSAAPAP 208
   : : : | | | : : : : : : : : : : : : : : : : : : : : : : : :
QY 131 IEKTIKAKQPREPQVYTLPPSRDELTKNQVSLTCLVKGFYPSDIAVEWESNGQPENNY 190
   : : : | | | : : : : : : : : : : : : : : : : : : : : : : : :
DB 209 IEKTIKGRKRAPQVYTIPTPKPEQMAKDKVSLTCLITDFFPEDITVEQWNGQPAENY 268
   : : : | | | : : : : : : : : : : : : : : : : : : : : : : : :
QY 191 KTTTPVLDSDGSFLLSKLTVDKSRWQOGNVFSCVMHEALHNHYTKSLSPGK 246
   : : : | | | : : : : : : : : : : : : : : : : : : : : : : : :
DB 269 KNTQPIMTNGSYFVYSKLVNQRKNWEAGNTFTCSVLHGLHNHTTKLSLSPGK 324

RESULT 12
GCC_RAT GCC_RAT STANDARD; PRT; 329 AA.
ID AC P20762;
DT 01-FEB-1991 (Rel. 17, Created)
DT 01-FEB-1991 (Rel. 17, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE I9 gamma-2C chain C region.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=88166903; PubMed=3127222;
RA Brueggemann M., Delmastro-Galfre P., Waldmann H., Calabi F.;
RT "Sequence of a rat immunoglobulin gamma 2c heavy chain constant
RT region cDNA: extensive homology to mouse gamma 3.";
RL Eur. J. Immunol. 18:317-319(1988).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; X07189; CAA30169.1; -
DR PIR; S00847; S00847.
DR HSSP; P01842; 7FAB.
DR InterPro; IPR003006; Ig_MHC.
DR InterPro; IPR003597; Ig-cl.
DR InterPro; IPR003600; Ig_like.
DR Pfam; PF00047; Ig; 2.
DR SMART; SM00410; Ig_like; 1.
DR SMART; SM00407; Igcl; 2.
DR PROSITE; PS00290; IG_MHC; 1.
KW Immunoglobulin domain; Immunoglobulin C region.
FT NON_TER 1 97
FT DOMAIN 1 97 CHI.
FT DOMAIN 98 113 HINGE.
FT DOMAIN 114 222 CH2.
FT DOMAIN 223 329 CH3.
FT DISULFID 15 15 INTERCHAIN (WITH A LIGHT CHAIN).
FT DISULFID 27 82
FT DISULFID 111 111 INTERCHAIN (WITH A HEAVY CHAIN).
FT DISULFID 113 113 INTERCHAIN (WITH A HEAVY CHAIN).
FT DISULFID 143 203
FT DISULFID 249 307
SQ SEQUENCE 329 AA; 36571 MW; 5FCD7B7933850773 CRC64;

```

```
Query Match 61.3%; Score 818; DB 1; Length 329;
Best Local Similarity 57.9%; Pred. No. 7.7e-21;
Matches 146; Conservative 42; Mismatches 38; Indels 26; Gaps 2;

QY 21 KTHTC-----PP-----CPAPELLGPGSVFLPPPKDITLMISR 54
DB 78 QTVCSTVAHPATKSNLIKRIEPRPKPRPTDICSDDNGLGRPSVFIFPPKPKDITLMTL 137
QY 55 TPEVTCVVVDVSHEDPEVKFNKYVDGVEVHNNAKTPREQYNSTYRVVSVLTVQLHODWLN 114
DB 138 TRKVCVVVDVSEEDPDVGFQSFVDNRVFTAQTPQHEQLNGTKRVVSTLHIQHDWMS 197
QY 115 GREYKCKVSNKALPAPIEKTISKAKGQPREPQVYTLPPSRDELTKNQVSLTCLVKGFGYPS 174
DB 198 GKEFKCKVNNKDLPSPIEKTISKAKGQPREPQVYTLPPSRDELTKNQVSLTCLVKGFGYPS 257
QY 175 DTAVFESNGQPNKYKTPPVLDSGSGFFLYSKLTVDKSRQOQGVNFSCSVNHEALNNH 234
DB 258 SISVWERNGELEDQYKNTLPVLDSDESFLYSLKSLVDTSWMRGDICTSVVHEALNNH 317
QY 235 YTKSLSLSPGK 246
DB 318 HTQKNLSRSPGK 329

RESULT 13
GC1M_MOUSE STANDARD; PRT; 393 AA.
AC P01869;
DT 21-JUL-1986 (Rel. 01, Created)
DT 01-AUG-1991 (Rel. 19, Last sequence update)
DT 30-MAY-2000 (Rel. 39, Last annotation update)
DE Ig gamma-1 chain C region, membrane-bound form.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=80045036; PubMed=115593;
RA Honjo T., Obata M., Yamawaki-Kataoka Y., Kataoka T., Kawakami T.,
RA Takahashi N., Mano Y.;
RT Cloning and complete nucleotide sequence of mouse immunoglobulin
RT gamma 1 chain gene.;
RL Cell 18:559-568(1979).
RN [2]
RP SEQUENCE OF 323-393 FROM N.A.
RX MEDLINE=82197626; PubMed=6804950;
RA Tyler B.M., Cowman A.F., Gerondakis S.D., Adams J.M., Bernard O.;
RT "mRNA for surface immunoglobulin gamma chains encodes a highly
RT conserved transmembrane sequence and a 28-residue intracellular
RT domain.";
RL Proc. Natl. Acad. Sci. U.S.A. 79:2008-2012(1982).
RN [3]
RP SEQUENCE OF 323-366 FROM N.A.
RX MEDLINE=82115295; PubMed=6799207;
RA Rogers J., Choi E., Souza L., Carter C., Word C.J., Kuehl M.,
RA Eisenberg D., Wall R.;
RT "Gene segments encoding transmembrane carboxyl termini of
RT immunoglobulin gamma chains.";
RL Cell 26:19-27(1981).
RN [4]
RP SEQUENCE OF 1-44 FROM N.A.
RX MEDLINE=8222190; PubMed=6283537;
RA Yamawaki-Kataoka Y., Nakai S., Miyata T., Honjo T.;
RT "Nucleotide sequences of gene segments encoding membrane domains of
RT immunoglobulin gamma chains.";
RL Proc. Natl. Acad. Sci. U.S.A. 79:2623-2627(1982).
CC -!- ALTERNATIVE PRODUCTS: CELL LINES PRODUCING IGG CONTAIN TWO MRNA
CC SPECIES FOR IG GAMMA CHAINS. THE MAJOR SPECIES ENCODES SECRETED
CC GAMMA CHAINS. A LESS ABUNDANT SPECIES APPEARS TO ENCODE MEMBRANE-
CC BOUND CHAINS IN THAT IT CONTAINS AN ALTERNATIVE 3' END, ENCODED
CC IN SEPARATE EXONS, THAT IS HOMOLOGOUS WITH THE MEMBRANE-BOUND

CC SEGMENT OF MU CHAINS.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; V00793; CAA24172.1; -
CC EMBL; V00793; CAA24173.1; -
CC EMBL; V00793; CAA24174.1; -
CC PIR; B02159; GLMSM.
CC HSP; P01842; 7FAB.
CC MGD; MGI:96446; Igh-4.
CC InterPro; IPR003006; Ig_MHC.
CC InterPro; IPR003597; Ig_c1.
CC Pfam; PF00047; Ig; 3.
CC SMART; SM00407; Igel; 2.
CC PROSITE; PS00290; IG_MHC; 1.
CC Immunoglobulin domain; Immunoglobulin C region; Glycoprotein;
CC Alternative splicing; Transmembrane.
CC NON_TER 1
CC DOMAIN 1 97
CC DOMAIN 98 110
CC DOMAIN 111 217
CC DOMAIN 218 324
CC DISULFID 27 82
CC DISULFID 102 102
CC DISULFID 104 104
CC DISULFID 107 107
CC DISULFID 109 109
CC DISULFID 138 198
CC CARBOHYD 174 174
CC DISULFID 244 302
CC TRANSMEM 340 357
CC DOMAIN 358 493
CC SEQUENCE 393 AA; 43386 MW; 4CC88343B7A1CE27 CRC64;

Query Match 61.3%; Score 818; DB 1; Length 393;
Best Local Similarity 59.6%; Pred. No. 1.3e-20;
Matches 140; Conservative 48; Mismatches 42; Indels 5; Gaps 2;

QY 13 KLKEAFMDKTKTCTPP--CPAPELLGPGSVFLPPPKDITLMISRTPETVCVVVDVSHBDP 70
DB 92 KVDKKIVPRDCGCKPCICTVEV---SSVFIFPPKPKDVLITITPKVTCVVVDISKDDP 148
QY 71 EVKFNWYVDGVEVHNNAKTPREQYNSTYRVVSVLTVQLHODWLNKGYKCKVSNKALPAP 130
DB 149 EVQSFWEFVDDVEVHTAQTPREQFNSTFRSVSELPIMHQDLNKGKFKCRVNSAAPPAP 208
QY 131 IEKTIKRAKGPQREPOVYTLPPSRDELTKNQVSLTCLVKGFGYPSDIAVWESNGQPENNY 190
DB 209 IEKTIKTRKGRKAPQVYTIPTPKQMAKDKVSLTCTMTDFFEDITVWQMNGQPAENY 268
QY 191 KTTTPPVLDGSGFFLYSKLTVDKSRQOQGVNFSCSVNHEALNNHYTKSLSLSPG 245
DB 269 KNTQPIINTNGSYFYSLKLVQKSNWEAGNTFTCSVLHLEGLNHHHTKSLSHSPG 323

RESULT 14
GCAA_MOUSE STANDARD; PRT; 330 AA.
ID GCAA_MOUSE
AC P01863;
DT 21-JUL-1986 (Rel. 01, Created)
DT 21-JUL-1986 (Rel. 01, Last sequence update)
DT 30-MAY-2000 (Rel. 39, Last annotation update)
DE Ig gamma-2A chain C region, A allele.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
```

1] SEQUENCE FROM N.A.  
MEDLINE-81076554; PubMed-6777755;  
Sikotav J., L., Auffray C., Rougeon F.;  
"Structure of the constant and 3' untranslated regions of the murine  
Balb/c gamma 2a heavy chain messenger RNA.";  
Nucleic Acids Res. 8:3143-3155(1980).  
[2]  
SEQUENCE FROM N.A.  
MEDLINE-81198976; PubMed-6262729;  
Yamawaki-Kataoka Y., Miyata T., Honjo T.;  
"The complete nucleotide sequence of mouse immunoglobulin gamma 2a gene  
and evolution of heavy chain genes: further evidence for intervening  
sequence-mediated domain transfer.";  
Nucleic Acids Res. 9:1365-1381(1981).  
[3]  
SEQUENCE FROM N.A.  
MEDLINE-81223894; PubMed-6787604;  
Olio R., Auffray C., Mochamps C., Rougeon F.;  
"Comparison of mouse immunoglobulin gamma 2a and gamma 2b chain genes  
family.";  
Proc. Natl. Acad. Sci. U.S.A. 78:2442-2446(1981).  
[4]  
MYELOMA PROTEIN MOPC 173.  
MEDLINE-74175517; PubMed-4831970;  
Bourgeois A., Fougereau M., Rocca-Serra J.;  
"Determination of the primary structure of a mouse IgG2a  
immunoglobulin: amino-acid sequence of the Fc fragment. Implications  
for the evolution of immunoglobulin structure and function.";  
Eur. J. Biochem. 43:423-435(1974).  
[5]  
DISULFIDE BONDS.  
MEDLINE-73056887; PubMed-4565406;  
de Preval C., Fougereau M.;  
"Determination of the primary structure of a mouse gamma G2a  
immunoglobulin. Identification of the disulfide bridges.";  
Eur. J. Biochem. 30:452-462(1972).  
-----  
This SWISS-PROT entry is copyright. It is produced through a collaboration  
between the Swiss Institute of Bioinformatics and the EMBL outstation -  
the European Bioinformatics Institute. There are no restrictions on its  
use by non-profit institutions as long as its content is in no way  
modified and this statement is not removed. Usage by and for commercial  
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
-----  
EMBL; V00798; CAA24178.1; -;  
PIR; A02152; G2MSA.  
HSSP; P01842; 7FAB.  
InterPro: IPR003006; Ig\_MHC.  
InterPro: IPR003597; Ig\_cl.  
InterPro: IPR003600; Ig\_like.  
Pfam; PF00047; Ig; 2.  
SMART; SM00410; Ig\_like; 1.  
SMART; SM00407; IGC1; 2.  
PROSITE; PS00290; IG\_MHC; 1.  
Immunoglobulin domain; Immunoglobulin C region.  
NON\_TER 1  
FT DISULFID 15 15 INTERCHAIN (WITH A LIGHT CHAIN).  
FT DISULFID 27 82  
FT DISULFID 107 107 INTERCHAIN (WITH A HEAVY CHAIN).  
FT DISULFID 110 110 INTERCHAIN (WITH A HEAVY CHAIN).  
FT DISULFID 112 112 INTERCHAIN (WITH A HEAVY CHAIN).  
FT DISULFID 144 204  
FT DISULFID 250 308  
FT MOD\_RES 330 330 REMOVED POST-TRANSLATIONALLY.  
SQ SEQUENCE 330 AA; 36389 MW; B84361C545A6864 CRC64;  
  
Query Match 61.0%; Score 814; DB 1; Length 330;  
Best Local Similarity 60.7%; Pred. No. 1.le-20;  
Matches 150; Conservative 31; Mismatches 56; Indels 10; Gaps 2;

QY 10 VAEKLKFAFMK-----THTCPP--CPAPELLGSPVFLFPKPKDTLMISRTPEVT 59  
Db 84 VAHPASSTKVKKIEPRGPTIKPCPPCKCAPNLLGGPSVFIFPKIKDVLMSLSPIVT 143  
QY 60 CVVVDVSHEDPEVKFNKYNVDGVEVHNNAKTPREBOYNSTYRVVSVLTVLHODWLNQKEYK 119  
Db 144 CVVVDVSEDDPDVOISWVFNNEVHTAQOTQTHREDYINSTLRVVSALPIQIHOHDMWSGKEFK 203  
QY 120 KYSNKKALPAPIETKISKARGQPREPOVYTLPPSRDELTKNQSLTCLVKGFYPSDIAVE 179  
Db 204 KYNKNDLPAPIETISKPGSVRAPOVYVLPPEEEMTKQVTLTDMVDFMEDIVYE 263  
QY 180 WESNGQENNYKTPPVLDSDGSGFLYSLKLTVDKSRQOQNVFSCSVHHEALHHNYTKS 239  
Db 264 WTNGKTELNYKTEPVLDSDGSGYFMYSKLRVEKKNNVERNSYSCSVVHHEGLHHHTKS 323  
  
QY 240 LSLSPGK 246  
Db 324 FSRTPGK 330  
  
RESULT 15  
GCAB\_MOUSE STANDARD; PRT; 335 AA.  
ID GCAB\_MOUSE AC P01864;  
DT 21-JUL-1986 (Rel. 01, Created)  
DT 21-JUL-1986 (Rel. 01, Last sequence update)  
DT 15-JUL-1999 (Rel. 38, Last annotation update)  
DE Ig gamma-2A chain C region, B allele.  
OS Mus musculus (Mouse).  
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
OX NCBI\_TaxID=10090;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=C57BL/6;  
RX MEDLINE-82037861; PubMed-6170065;  
RA Schreier P.H., Bothwell A.L.M., Mueller-Hill B., Baltimore D.;  
RT "Multiple differences between the nucleic acid sequences of the  
IgG2a and IgG2b alleles of the mouse.";  
Proc. Natl. Acad. Sci. U.S.A. 78:4495-4499(1981).  
RN [2]  
RP SEQUENCE.  
RX MEDLINE-82037777; PubMed-6794027;  
RA Dognin M.J., Lauwereys M., Strosberg A.D.;  
RT "Multiple amino acid substitutions between murine gamma 2a heavy  
chain Fc regions of IgA and IgB allotypic forms.";  
Proc. Natl. Acad. Sci. U.S.A. 78:4031-4035(1981).  
RL -1- MISCELLANEOUS: THE SEQUENCE DIFFERS FROM THAT OF THE ALLELE,  
CC FROM BALB/C MICE, AT 15% OF THE POSITIONS.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
between the Swiss Institute of Bioinformatics and the EMBL outstation -  
the European Bioinformatics Institute. There are no restrictions on its  
use by non-profit institutions as long as its content is in no way  
modified and this statement is not removed. Usage by and for commercial  
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
CC EMBL; J00479; -; NOT\_ANNOTATED\_CDS.  
DR PIR; A02153; G2MSAB.  
DR HSSP; P01842; 7FAB.  
DR InterPro: IPR003006; Ig\_MHC.  
DR InterPro: IPR003597; Ig\_cl.  
DR InterPro: IPR003600; Ig\_like.  
DR Pfam; PF00047; Ig; 3.  
DR SMART; SM00410; IG\_like; 1.  
DR SMART; SM00407; IGC1; 2.  
DR PROSITE; PS00290; IG\_MHC; 1.  
DR Immunoglobulin domain; Immunoglobulin C region.  
NON\_TER 1  
FT SEQUENCE 335 AA; 36596 MW; FA3382792CBB13C6 CRC64;  
SQ

Query Match 60.6%; Score 809; DB 1; Length 335;  
Best Local Similarity 55.8%; Pred. No. 1.7e-20;  
Matches 144; Conservative 39; Mismatches 43; Indels 32; Gaps 1;

QY	21	KTHTC-----	PPCPAPELLGGPSVFLPPKPKD	48
DB	78	QTITCNVAHPASTKVYDKIEPRVPIQNPCPPHQRVPPCAAPDLLGGPSVFIFPPKIKD	137	
QY	49	TLMISRTPEVTCVVVDVSHEDPEVKFNWYVDGVEVHNAKTKPREQYNSTYRVVSVLT	108	
DB	138	VLMSISLSPMWTCTVVVDVSEDDVDQISWVFNVEVHTAQTQTHREDYNSTLRVVSALPIQ	197	
QY	109	HODWLNKEYCKVSNKALPAPIEKTISKAKGQPREPQVYTLPPSRDELTKNQVSLTCLV	168	
DB	198	HODWMSGKEFKCKVNNRALPSPIEKTISKPRGPRVAPQVYVLPPEAEEMTKKEFSLT	257	
QY	169	KGFYPSDIAVENESNGOPENNYKTTTPVLDSDGSFELYSKLTVDKSRWQOQNVFSCSV	228	
DB	258	TGFLPAEIAVDWTSNGRTQNYKNTATVLDSDGSYFMYSKLRVQKSTWERSL	317	
QY	229	EALHNNHYTKSLSLSPGK	246	
DB	318	EVLNHLTTTKTISRSLGK	335	

Search completed: April 21, 2003, 10:43:11  
Job time : 10.4059 secs

GenCore version 5.1.4.p5.4578  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: April 21, 2003, 10:42:21 ; Search time 36.1765 seconds  
(without alignments)  
1401.120 million cell updates/sec

Title: 7LINK2

Perfect score: 1334

Sequence: 1 DWLKAFYDKVAEKLKEAFMD.....MHEALHNHYTKQSLSLSPCK 246

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.0

Searched: 671580 seqs, 206047115 residues

Total number of hits satisfying chosen parameters: 671580

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

SPREMBL\_21.\*  
1: sp\_archaea.\*  
2: sp\_bacteria.\*  
3: sp\_fungi.\*  
4: sp\_human.\*  
5: sp\_invertebrate.\*  
6: sp\_mammal.\*  
7: sp\_mhc.\*  
8: sp\_organelle.\*  
9: sp\_phase.\*  
10: sp\_plant.\*  
11: sp\_rodent.\*  
12: sp\_virus.\*  
13: sp\_vertebrate.\*  
14: sp\_unclassified.\*  
15: sp\_rvirus.\*  
16: sp\_bacteriap.\*  
17: sp\_archaeap.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1250	93.7	471	4	O8TC77
2	1242	93.1	701	4	Q96PQ8
3	1151	86.3	473	4	O8TC63
4	917	68.7	337	6	Q95M34
5	847	63.5	463	11	Q99LC4
6	843	63.2	469	11	O8R3V9
7	838	62.8	437	11	O9R1A4
8	829	62.1	473	11	Q9D8L4
9	824	61.8	473	11	Q99L25
10	821	61.5	468	11	Q99L31
11	797	59.7	474	11	O8R3H6
12	794	59.5	473	11	O91Z05
13	486	36.4	34350	4	O8WZ42
14	463	34.7	26926	4	Q10466
15	463	34.7	26926	4	O8WZB3
16	406	30.4	16215	5	Q9NFS3

17	405	30.4	17352	5	Q95YM2
18	392	29.4	597	4	Q96BB9
19	390	29.2	7962	4	O10465
20	387	29.0	597	4	Q9BQB8
21	387	29.0	597	4	Q9BU10
22	387	29.0	614	4	Q96GA6
23	379	28.4	375	4	Q9BSZ1
24	377	28.3	588	4	O8WUX4
25	377	28.3	613	4	O8WUK1
26	377	28.3	618	4	Q96AA6
27	375	28.1	613	4	Q96EY0
28	372	27.9	613	11	Q8VCX7
29	371	27.8	15281	3	Q09164
30	358	26.8	6632	5	O01761
31	357	26.8	6632	5	O17362
32	356	26.7	7107	5	Q9V4F7
33	352	26.4	5636	4	Q96RW7
34	347	26.0	5198	5	O76518
35	347	26.0	5604	4	O8WZ53
36	345	25.9	6831	5	O23550
37	345	25.9	7160	5	O23551
38	344	25.8	8563	2	Q54297
39	343	25.7	4824	5	Q95YM1
40	342	25.6	6658	5	O76281
41	342	25.6	6875	6	O28733
42	340	25.5	8805	5	O9V6V4
43	339	25.4	13055	5	Q09165
44	337	25.3	9376	2	O85168
45	331	24.8	7463	16	Q9Z4X6

## ALIGNMENTS

RESULT 1

Q8TC77 ID Q8TC77 PRELIMINARY; PRT; 471 AA.  
AC O8TC77;  
DT 01-JUN-2002 (TREMBLrel. 21, Created)  
DT 01-JUN-2002 (TREMBLrel. 21, Last sequence update)  
DE Hypothetical 51.8 kDa protein.  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
OX NCBI\_TaxID=9606;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC TISSUE=SPLEEN;  
RA Strausberg R.;  
RL Submitted (FEB-2002) to the EMBL/GenBank/DBJ databases.  
DR EMBL; BC024289; A024289.1;  
KW Hypothetical protein.  
SQ SEQUENCE 471 AA; 51791 MW; 388F7F4CF588660E CRC64;

Query Match 93.7%; Score 1250; DB 4; Length 471;  
Best Local Similarity 57.0%; Pred. No. 7.6e-44;  
Matches 232; Conservative 8; Mismatches 6; Indels 161; Gaps 2;  
QY 1 DWLKA-----  
Db 65 EWVSSMSSSSYYVADSVKGRFTISRDNKAKSLYLQMSLRARDTAVVYCARDLRQLTS 124  
QY 6 --FYD-----  
Db 125 YWYFDLMGRGTLVTSSASTKGPSVFFLAPSSKTSGGTAALGCLVKDYPEPVTYVSWNS 184  
QY 9 -----KVAEKLKEAFM 19  
Db 185 GALTSGVHTTFAVLQSSGLYSLSSVTVPSSSLGTQTYICNVNKPNTKVDKKEPKSC 244  
QY 20 DKHTCTPCPAPELLGGPSVFLFPPKPKDTLMISRTPEVTCVVVDVSHEDPEVKFNWYVD 79  
|||||

```
Db 245 DKHTCPCPAPPELLGGPSVFLFPPPKPDTLMSRTPEVTCVVVDVSHEDPEVKFNWYVD 304
QY 80 GVEVHNAKTKPREQNSYRVSVLTLVHQLDNLNGKEYCKVSNKALPAPIETKTSKAK 139
Db 305 GVEVHNAKTKPREQNSYRVSVLTLVHQLDNLNGKEYCKVSNKALPAPIETKTSKAK 364
QY 140 GQREPOVYTLPSRDELTKNQVSLTCLVKGFYPSDIAVEWESNGQPENNYKTTPPVLD 199
Db 365 GQREPOVYTLPSRDELTKNQVSLTCLVKGFYPSDIAVEWESNGQPENNYKTTPPVLD 424
QY 200 DGSFFLYSKLTVDKSRWQGNFSCSVMHAEALHNYHTQKLSLSPGK 246
Db 425 DGSFFLYSKLTVDKSRWQGNFSCSVMHAEALHNYHTQKLSLSPGK 471

RESULT 2
Q96PQ8
ID Q96PQ8 PRELIMINARY; PRT: 701 AA.
AC Q96PQ8;
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-WAR-2002 (TrEMBLrel. 20, Last annotation update)
DE Factor VII active site mutant immunocjugate.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=21477448; PubMed=11593034;
RA Hu Z., Garen A.;
RT "Targeting tissue factor on tumor vascular endothelial cells and tumor
RT cells for immunotherapy in mouse models of prostatic cancer.";
RL Proc. Natl. Acad. Sci. U.S.A. 98:12180-12185(2001).
DR EMBL: AF272774; AAK58686.1; -.
DR InterPro: IPR000152; Asx_hydroxyl.
DR InterPro: IPR000561; EGF-like.
DR InterPro: IPR000742; EGF 2.
DR InterPro: IPR001881; EGF Ca.
DR InterPro: IPR003006; Ig_MHC.
DR InterPro: IPR001254; Ser_protease_Try.
DR InterPro: IPR000294; Vitk_dep_GLA.
DR Pfam: PF00008; EGF; 2.
DR Pfam: PF00594; gla; 1.
DR Pfam: PF00047; ig; 2.
DR Pfam: PF00089; trypsin; 1.
DR SMART: SM00181; EGF; 2.
DR PROSITE: PS00010; ASX_HYDROXYL; UNKNOWN_1.
DR PROSITE: PS00022; EGF_1; UNKNOWN_1.
DR PROSITE: PS01186; EGF 2; UNKNOWN_1.
DR PROSITE: PS01187; EGF Ca; UNKNOWN_1.
DR PROSITE: PS00011; GLU CARBOXYLATION; UNKNOWN_1.
DR PROSITE: PS00290; IG_MHC; UNKNOWN_1.
DR PROSITE: PS00240; TRYPSIN_DOM; 1.
DR PROSITE: PS00134; TRYPSIN_HIS; UNKNOWN_1.
DR PROSITE: PS00135; TRYPSIN_SER; UNKNOWN_1.
KW Hydrolase; Serine protease.
SQ SEQUENCE 701 AA; 77826 MW; 94AC6CEB42CC992F CRC64;

Query Match 93.1%; Score 1242; DB 4; Length 701;
Best Local Similarity 87.2%; Pred. No. 6.5e-43;
Matches 231; Conservative 3; Mismatches 4; Indels 27; Gaps 2;

QY 1 DMLKAFYDKVAELKEA-----FMDKTHTCPPAPPELLGGPSVFL 41
Db 445 EWL-----QKLMRPRPGLVLRAPFGSAEPKSCDKTHTCPPAPPELLGGPSVFL 496

QY 42 FPPKPKDTLMISRTPEVTCVVVDVSHEDPEVKFNWYVDGVEVHNAKTKPREEQYNSTYRV 101
Db 497 FPPKPKDTLMISRTPEVTCVVVDVSHEDPEVKFNWYVDGVEVHNAKTKPREEQYNSTYRV 556
QY 102 VSVLTVLHODWLNKGEYCKVSNKALPAPIETKTSKAKQPREPVYTLPPSRDELTKNQ 161
```

```
Db 557 VSVLTVLHODWLNKGEYCKVSNKALPAPIETKTSKAKQPREPVYTLPPSRDELTKNQ 616
QY 162 VSLTCLVKGFYPSDIAVEWESNGQPENNYKTTPPVLDSDGSFPLYSKLTVDKSRWQGNV 221
Db 617 VSLTCLVKGFYPSDIAVEWESNGQPENNYKTTPPVLDSDGSFPLYSKLTVDKSRWQGNV 676
QY 222 FSCSVMHAEALHNYHTQKLSLSPGK 246
Db 677 FSCSVMHAEALHNYHTQKLSLSPGK 701

RESULT 3
Q8TC63
ID Q8TC63 PRELIMINARY; PRT: 473 AA.
AC Q8TC63;
DT 01-JUN-2002 (TrEMBLrel. 21, Created)
DT 01-JUN-2002 (TrEMBLrel. 21, Last sequence update)
DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE Hypothetical 52.0 kDa protein.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=KIDNEY;
RA Strausberg R.;
RL Submitted (MAR-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL: BC025985; AAH25985.1; -.
KW Hypothetical protein.
SQ SEQUENCE 473 AA; 51986 MW; E29920B09BA369F5 CRC64;

Query Match 86.3%; Score 1151; DB 4; Length 473;
Best Local Similarity 54.2%; Pred. No. 1.2e-39;
Matches 217; Conservative 16; Mismatches 13; Indels 154; Gaps 4;

QY 1 DML-----KAFYDKVAE----- 12
Db 74 EWIGTINFGNMYSPSLRSRVMTSADMSKNSFYLKLDSTVTAADTAVYYCAAGHLVMGFG 133
QY 13 -----KLUKAPMD----- 20
Db 134 AHWGQGLVSPASTKGPSVFPLAPCSKSTSESTAALGCLVKDYFPEPTVYSWNSGALT 193
QY 21 -----KTHF----- 26
Db 194 SGVHTFPAVLQSSGLYSLSSVVTVPSSSLGTQYTCNVNDHKSNTYVKDKRVESKYGPPCP 253
QY 27 PCPAPPELLGGPSVFLFPPPKPDTLMSRTPEVTCVVVDVSHEDPEVKFNWYVDGVEVHNA 86
Db 254 SCPAPEELGGPSVFLFPPPKPDTLMSRTPEVTCVVVDVSHEDPEVKFNWYVDGVEVHNA 313
QY 87 KTKPREEQYNSTYRVVSVLTVHLQDWLNKGEYCKVSNKALPAPIETKTSKAKQPREPQ 146
Db 314 KTKPREEQYNSTYRVVSVLTVHLQDWLNKGEYCKVSNKALPAPIETKTSKAKQPREPQ 373
QY 147 VYTLPPSRDELTKNQVSLTCLVKGFYPSDIAVEWESNGQPENNYKTTPPVLDSDGSF 206
Db 374 VYTLPPSRDELTKNQVSLTCLVKGFYPSDIAVEWESNGQPENNYKTTPPVLDSDGSF 433
QY 207 SKLTVDKSRWQGNFSCSVMHAEALHNYHTQKLSLSPGK 246
Db 434 SRLTVDKSRWQGNFSCSVMHAEALHNYHTQKLSLSPGK 473

RESULT 4
Q95M34
ID Q95M34 PRELIMINARY; PRT: 337 AA.
AC Q95M34;
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE Immunoglobulin gamma 1 heavy chain constant region
```

```

DE (Fragment).
GN IGH1.
OS Equus caballus (Horse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Perissodactyla; Equidae; Equus.
OX NCBI_TaxID=9796;
RN [1]
RP SEQUENCE FROM N.A.
RA Wagner B.; (DEC-2000) to the EMBL/GenBank/DBJ databases.
RL Submitted (DEC-2000) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=98383416; PubMed=9717671;
RA Wagner B.; Overesch G.; Sheoran A.; Holmes M.; Richards C.;
RA Leibold W.; Radbruch A.;
RT "Organization of the equine immunoglobulin heavy chain constant region
RT genes. III. Alignment of c-mu, c-gamma, c-epsilon and c-alpha genes.";
RL Immunobiology 199;105-119(1998).
DR EMBL; AJ300675; CAC44624.1;
DR InterPro; IPR003006; Ig_MHC.
DR Pfam; PF00047; Ig; 2.
DR PROSITE; PS00290; IG_MHC; UNKNOWN_2.
FT NON_TER
SQ SEQUENCE 337 AA; 37438 MW; A60BF2B01DEFD1F6 CRC64;

Query Match 68.7%; Score 917; DB 6; Length 337;
Best Local Similarity 63.8%; Pred. No. 3.4e-30;
Matches 169; Conservative 33; Mismatches 40; Indels 23; Gaps 4;

QY 2 WLKAFY-----DKVAEKL-----KEAFMDKTHTCPCPAPELLGGPSVFLFP 43
DB 76 WTSETYICNVVHAASNEKFKVRIPIPDNHQKVCDSK---CPKCPAPELGGPSVIFP 132
QY 44 PKPKDTLMIKRTPEVTCVVDVSHEDPEVKFNWYDGVGVHNAKTKPREQYNSTYRV 103
DB 133 PNPKDTLMIKRTPEVTCVVDVSHEDPEVKFNWYDGVGVHNAKTKPREQYNSTYRV 192
QY 104 VLTVLHODWLNGLKGYCKVSNKALPAPIETKISAKQPREQVYTLPPSRDELTKNQV 163
DB 193 VLRIOHODWLNGLKGYCKVSNKALPAPIETKISAKQPREQVYTLPPSRDELTKNQV 252
QY 164 LTCLVKGFFSDIAVESNGQP--ENNYKTPPVLDSDGFFLYSKLTVDKSRWQGNV 221
DB 253 VTCLVKDFYPPFIEINWQSGQPELETYSTTQAQSDGSGFFLYSKLSVDRNRWQGGTT 312
QY 222 FSCSVMEALHNHYTQKSLSLSPCK 246
DB 313 FTGCVMEALHNHYTQKSNKPNCK 337

RESULT 5
Q99LC4 Q99LC4 PRELIMINARY; PRT; 463 AA.
AC Q99LC4; 2001 (TrEMBLrel. 17, Created)
DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT 01-JUN-2001 (TrEMBLrel. 21, Last annotation update)
DE Similar to RIKEN CDNA 181006O009 gene.
GN IGH-4.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RA Strausberg R.;
RL Submitted (FEB-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC003435; AAH03435.1;
DR HSSP; P01842; 7FAB.
DR MGD; MGI:96446; Igh-4.
DR InterPro; IPR003599; Ig.
DR InterPro; IPR003597; Ig-cl.
DR InterPro; IPR003600; Ig_like.

```

```

DR InterPro; IPR003006; Ig_MHC.
DR InterPro; IPR003596; Ig_v.
DR Pfam; PF00047; Ig; 4.
DR SMART; SM00409; IG; 2.
DR SMART; SM00407; IGcl; 2.
DR SMART; SM00406; IGV; 1.
DR SMART; SM00410; IG_Like; 1.
DR PROSITE; PS00290; IG_MHC; UNKNOWN_1.
SQ SEQUENCE 463 AA; 51007 MW; EAA674C6BBC30783 CRC64;

Query Match 63.5%; Score 847; DB 11; Length 463;
Best Local Similarity 37.1%; Pred. No. 9.5e-27;
Matches 154; Conservative 46; Mismatches 39; Indels 176; Gaps 8;

QY 2 WLK-----AFYDKVAEKLK-----EAFMD----- 20
DB 55 WKQRTQGLEWGEIYVPGSGNTYY---SEKFKGKATLTDDKSSSTAYMHLSSLTSEDSA 111
QY 21 -----KT----- 22
DB 112 VYFCARSSYSYDLFAYWGQGLTVTVSAAKTTPSVYVPLAPGSAQTNSMVTGLGLVKGY 171
QY 23 -----HT----- 24
DB 172 FPEPVTVMNSGSLSSGVHTFFPAVLQSDLYTLSSSVTVPSSTWPSSETVTCNVAHPASSTK 231
QY 25 -----CPD--CPAPELGGPSVFLFPKPKDTLMSRTPETVCVVVDVSHEDPE 71
DB 232 VDKKIVPRDGCRCPCICTVEV---SSVIFPPKPKDVLITLTPKVTVCVVVDISKDDPE 288
QY 72 VKFNWYVDGVGVHNAKTKPREQYNSTYRVVSVLTVLDHQLNGKEYCKVKSNKALPAPI 131
DB 289 VQFSWFVDVGVHNAKTKPREQYNSTYRVVSVLTVLDHQLNGKEYCKVKSNKALPAPI 348
QY 132 EKTISKAKGQPREQVYTLPPSRDELTKNQVSLTCLVKGFYPSDIAVESNGQPEPNYK 191
DB 349 EKTISKAKGQPREQVYTLPPSRDELTKNQVSLTCLVKGFYPSDIAVESNGQPEPNYK 408
QY 192 TTPPVLDSDGFFLYSKLTVDKSRWQGNVFCSCVMHEALHNHYTQKSLSLSPCK 246
DB 409 NTQPIMDTDSYFIYSLNVLQKSNWEAGNTFTCSVLHEGLHNHTKSLSHSPCK 463

RESULT 6
Q8R3V9 Q8R3V9 PRELIMINARY; PRT; 469 AA.
AC Q8R3V9;
DT 01-JUN-2002 (TrEMBLrel. 21, Created)
DT 01-JUN-2002 (TrEMBLrel. 21, Last sequence update)
DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE Hypothetical 52.0 kDa protein.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RA Strausberg R.;
RL Submitted (MAR-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC024405; AAH24405.1;
DR Hypothetical protein.
KW Hypothetical protein.
SQ SEQUENCE 469 AA; 51976 MW; 534793F155D05457 CRC64;

Query Match 63.2%; Score 843; DB 11; Length 469;
Best Local Similarity 36.8%; Pred. No. 1.5e-26;
Matches 150; Conservative 46; Mismatches 47; Indels 165; Gaps 5;

QY 1 DWL----- 3
DB 65 EWLGFIKANGYTYTEYSASVKGRFTISRDNSSQSLYLQMNALRAEDSATYICARDRSS 124
QY 4 -----KAFYDK----- 9

```

```

Db 125 YYSGTSFAYWGGTTLVTVSAAKTTPPSVYPLAPGSAQTNSMVTGLCLVKGYPPEPVTV 184
QY 10 -----VAEKLKEAFMDKTHT - 24
Db 185 TWNSGSLSSGVHTFPVAVLQSLDLYTLSSSVTVPSSTWPSVTCNVNHAHPASSTKVKKIYP 244
QY 25 ----CPP--CPAPELLGGPSVFLFPKPKDITLMISRTPEVTCVVVDVSHEDPEVKFNWTV 78
Db 245 RDCGCKPCTCTPEV---SSVFIFPPKPKDVLITITLTPKVTCCVVVDISKDDPEVQFSWFV 301
QY 79 DGEVHNATKPREEOYNTYRVVSVTLVHQQDLNGKEYCKVSKNALPAPIEKTISKA 138
Db 302 DDVEVHTAQTKPREOQNTFSRSELPIMHQDLNGKEFKRCVNSAFAPIEKTISKT 361
QY 139 KGPREPOVYTLPPSRDELTKNOVSLTCLVKGFPSPDIADVESNGOPENNYKTPPVLD 198
Db 362 KGRKAPQVYITPPRKEQAKDKVSLTCMTDFFPEDITVWQNGQPAENYKNTQPIMD 421
QY 199 SDGSFFLYSKLTVDKSRWQGNVFCSCVMHEALHNHYTKQSLSLSPGK 246
Db 422 TDGSYFVSKLVNOKSNWEAGNTFTCSVLHGLHNHTEKLSLHSPGK 469

RESULT 7
Q9R1A4
ID Q9R1A4 PRELIMINARY; PRT; 437 AA.
AC Q9R1A4;
DT 01-MAY-2000 (TREMBlrel. 13, Created)
DT 01-MAY-2000 (TREMBlrel. 13, Last sequence update)
DE 01-JUN-2002 (TREMBlrel. 21, Last annotation update)
GN Gamma1 heavy chain of Mab7 (Fragment).
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RA Wilde K.G., Yu X., Ekramoddoullah A.K.M., Misra S.;
RT "Cloning of cDNAs encoding for anti-white pine blister rust monoclonal
RT antibody (Mab 7, its light and heavy chains) and construction of a
RL single chain antibody (scFv).";
RL Submitted (MAY-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF152372; AAD40243.1; -.
DR HSSP; P01842; 7FAB.
DR MGD; MGI:96446; Igh-4.
DR InterPro; IPR003600; Igh-like.
DR InterPro; IPR003006; Igh_MHC.
DR InterPro; IPR003596; Igh_V.
DR Pfam; PF00047; Igh; 4.
DR SMART; SM00406; IGV; 1.
DR SMART; SM00410; IG_Like; 2.
DR PROSITE; PS00290; IG_MHC; UNKNOWN_1.
FT NON_TER 1
FT NON_TER 437
SQ SEQUENCE 437 AA; 48142 MW; 5C3A7BB3EE7D697C CRC64;

Query Match 62.8%; Score 838; DB 11; Length 437;
Best Local Similarity 37.4%; Pred. No. 1.9e-26;
Matches 148; Conservative 48; Mismatches 47; Indels 153; Gaps 5;

QY 1 DWLKAF-----YDKVAEKLKEAFMDK----- 21
Db 45 ENVAFSGGIYITDTSVKGRFTYKDKDRNLTSLQMSLSRSDTAMYYCARGDYSAYWG 104
QY 22 -----TH 23
Db 105 PGLTVTVSAAKTTPPSVYPLAPGSAQTNSMVTGLCLVKGYPPEPVTWNSGSLSSGVH 164
QY 24 T-----CPB--CPA 30
Db 165 TFPVAVLQSLDLYTLSSSVTVPSSTWPSVTCNVNHAHPASSTKVKKIYPRDCCCKPCICTV 224

```

```

QY 31 PELLGGPSVFLFPKPKDITLMISRTPEVTCVVVDVSHEDPEVKFNWYDGVSHNAKTP 90
Db 225 PEV---SSVFIFPPKPKDVLITITLTPKVTCCVVVDISKDDPEVQFSWFVDDVEVHTAQTP 281
QY 91 REEQNSTYRVVSVTLVHQQDLNGKEYCKVSKNALPAPIEKTISKAKGQPREPOVYTL 150
Db 282 REEQNSTFRSSELPIMHQDLNGKEFKRCVNSAFAPIEKTISKTKGRKPAQVYTI 341
QY 151 PPSRDELTKNOVSLTCLVKGFPSPDIADVESNGOPENNYKTPPVLDSDGSFFLYSKLT 210
Db 342 PPKPKAMKDKVSLTCMTDFFPEDITVWQNGQPAENYKNTQPIMDTDGSFYFVSKLN 401
QY 211 VDKSRWQGNVFCSCVMHEALHNHYTKQSLSLSPGK 246
Db 402 VQKSNWEAGNTFTCSVLHGLHNHTEKLSHSPGK 437

RESULT 8
Q9D8L4
ID Q9D8L4 PRELIMINARY; PRT; 473 AA.
AC Q9D8L4;
DT 01-JUN-2001 (TREMBlrel. 17, Created)
DT 01-JUN-2001 (TREMBlrel. 17, Last sequence update)
DE 01-DEC-2001 (TREMBlrel. 19, Last annotation update)
DE 181006009Rik protein.
GN IGH-1 OR 181006009RIK.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RA STRAIN=C57BL/6J; TISSUE=PANCREAS;
RA MEDLIN=21085660; PubMed=11217851;
RA Kawai J., Shinagawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,
RA Arakawa T., Hara A., Fukunishi Y., Konno H., Adachi J., Fukuda S.,
RA Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yananaka I.,
RA Saito T., Okazaki Y., Gojobori T., Bono H., Kasukawa T., Saito R.,
RA Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,
RA Fleischmann W., Gaasterland T., Gliss C., King B., Kochiwa H.,
RA Kuehl P., Lewis S., Matsuo Y., Nikaido I., Pesole G., Quackenbush J.,
RA Schriml L.M., Staebli F., Suzuki R., Tomita M., Wagner L., Washio T.,
RA Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barsh G.,
RA Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaudo M.F.,
RA Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.,
RA Gustincich S., Hill D., Hofmann M., Hume D.A., Kamiya M., Lee N.H.,
RA Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombauer P.,
RA Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,
RA Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F.,
RA Suzuki H., Toyooka K., Wang K.H., Weitz T., Whittaker C., Wilming L.,
RA Wynshaw-Boris A., Yoshida K., Hasegawa Y., Kawaji H., Kohtsuki S.,
RA Hayashizaki Y.;
RT "Functional annotation of a full-length mouse cDNA collection.";
RL Nature 409:685-690(2001).
DR EMBL; AK007918; BAB25349.1; -.
DR HSSP; P01842; 7FAB.
DR MGD; MGI:96443; Igh-1.
DR InterPro; IPR003599; Igh.
DR InterPro; IPR003597; Igh-cl.
DR InterPro; IPR003600; Igh-like.
DR InterPro; IPR003006; Igh_MHC.
DR InterPro; IPR003596; Igh_V.
DR Pfam; PF00047; Igh; 4.
DR SMART; SM00409; IG; 2.
DR SMART; SM00407; IGV; 3.
DR SMART; SM00406; IGV; 1.
DR PROSITE; PS00290; IG_MHC; UNKNOWN_1.
SQ SEQUENCE 473 AA; 51699 MW; 9DED57A514475FBB CRC64;

Query Match 62.1%; Score 829; DB 11; Length 473;
Best Local Similarity 36.4%; Pred. No. 5.9e-26;
Matches 154; Conservative 43; Mismatches 46; Indels 180; Gaps 6;

```



```
QY 1 DWLK-----AFYDKVAEKLK-----EAFMD----- 20
Db 54 NWKQRPQGLEWIGKIGPGSGSTYIN---EKFKGATLTADKSSSTAYMQLSLTSEDS 110
QY 21 -----KT----- 22
Db 111 AVYFCARSGYDYDFAYWYGQGLTWTVSAAKTTAPSVVPLAPVCGGTGSSVTGLCLVKGY 170
QY 23 -----HT----- 24
Db 171 FPEPVLTNWSSLSGSLGVTFFPALLOGLYTLSSSVTVTSNTWPSQITITCNVAHPASSTK 230
QY 25 -----CPPCAPPELLGGPSVFLFPKPKDITLMISRTPEVTCVVV 63
Db 231 VDKKIEPRVITQPCPPLEKPCPCAPDLGGPSVFIFFPKIKDVLMSLSPMVTCCVV 290
QY 64 DVSHEDPEVFNKVVYDVEVHNNAKTRPREQYNSTRVSVLVHLDWLNKGEYKCKVS 123
Db 291 DVSEDDPDVQISFVNNVEVHTAQOTQTHREDYNSTLRVVSALPIQHDWMSGKEFKCKVN 350
QY 124 NKALPAPIETISKAKQPREPOVYITLPPSRDELTKNOVSLTCLVKGFPYSDIAVEWESN 183
Db 351 NRALPSPIETISKPRGPVAPQYVILPPPAEEMTKKEFSLTCTMTGFLPAEIAVDWTSN 410
QY 184 GOPENNTKTTPPVLDSDGSFSLXSKLTVDKSRWQOQNVFSCSVMHALHNHYTKQSLSL 243
Db 411 GRTEQNYKNTATVLDSDGSYFMSKLRVQKSTWERSLFCACSVVHGLHNHLLTKTISRS 470
QY 244 PGK 246
Db 471 LGK 473

RESULT 9
Q99L25 PRELIMINARY; PRT; 473 AA.
AC Q99L25;
DT 01-JUN-2001 (TrEMBLrel. 17, Created)
DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE Similar to RIKEN CDNA 1810060009 gene.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RA Strausberg R.;
RL Submitted (FEB-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC003888; AA03888.1; -
DR HSSP; P01842; 7FAB.
DR InterPro; IPR003599; Ig.
DR InterPro; IPR003597; Ig-cl.
DR InterPro; IPR003600; Ig-like.
DR InterPro; IPR003006; Ig_MHC.
DR Pfam; PF00047; Ig; 3.
DR SMART; SM00409; IG; 2.
DR SMART; SM00407; IGC1; 3.
DR SMART; SM00406; IGV; 1.
DR SMART; SM00410; IG_like; 1.
DR PROSITE; PS00290; IG_MHC; UNKNOWN_1.
SQ SEQUENCE 473 AA; 52449 MW; BE9889B7986DA155 CRC64;

Query Match 61.8%; Score 824; DB 11; Length 473;
Best Local Similarity 38.6%; Pred. No. 9.7e-26;
Matches 158; Conservative 33; Mismatches 55; Indels 163; Gaps 6;

QY 1 DWLKAFY-----DKVAEKLK-----EAFM----- 19
Db 65 EWIGIYIPRGSTKYNEKFKGATLTADKSSSTAYMQLNSLTSEDSAVCFCSRGGSYYG 124
```

```
QY 20 -----DKT----- 22
Db 125 YGLYFDYWGQGTITVSSAKTTAPSVVPLAPVCGDTTGSSTVLGCLVKGYFPEPVTLTW 184
QY 23 -----HT----- 24
Db 185 NSGSLSGVHTFFPAVLQSDLYTLSSSVTVTSSTWPSQSITCNVAHPASSTKVDKKIEPRG 244
QY 25 -----CPP--CAPPELLGGPSVFLFPKPKDITLMISRTPEVTCVVVDVSHEDPEVKFNWY 77
Db 245 PTKPCPCPCAPNLLGGPSVFIFFPKIKDVLMSLSPMVTCCVVVDVSEDDPDVQLSWF 304
QY 78 VDGVEVHNNAKTRPREQYNSTRVSVLVHLDWLNKGEYKCKVSNKALPAPIETISK 137
Db 305 VNNVEVLTAQTQTHREDYNSTLRVVSALPIQHDWMSGKEFKCKVNNKALPAPIETISK 364
QY 138 AKQPREPOVYITLPPSRDELTKNOVSLTCLVKGFPYSDIAVEWESNQPNNTKTTPPVL 197
Db 365 PKGSVRAPQYVILPPPAEEMTKKQVTLTCTMTDFMPEDIYVEWTNNGKTELNYKNTPEVL 424
QY 198 DSDGSFSLXSKLTVDKSRWQOQNVFSCSVMHALHNHYTKQSLSLSPGK 246
Db 425 DSDGSYFMSKLRVEKKNWVERNYSYCSVVHGLHNHHTKTSFRTPECK 473

RESULT 10
Q99L31 PRELIMINARY; PRT; 468 AA.
AC Q99L31;
DT 01-JUN-2001 (TrEMBLrel. 17, Created)
DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE Similar to RIKEN CDNA 1810060009 gene.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RA Strausberg R.;
RL Submitted (FEB-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC003878; AA03878.1; -
DR HSSP; P01842; 7FAB.
DR InterPro; IPR003599; Ig.
DR InterPro; IPR003597; Ig-cl.
DR InterPro; IPR003600; Ig-like.
DR InterPro; IPR003006; Ig_MHC.
DR Pfam; PF00047; Ig; 3.
DR SMART; SM00409; IG; 2.
DR SMART; SM00407; IGC1; 3.
DR SMART; SM00406; IGV; 1.
DR SMART; SM00410; IG_like; 1.
DR PROSITE; PS00290; IG_MHC; UNKNOWN_1.
SQ SEQUENCE 468 AA; 51661 MW; 96352328B3332ADB CRC64;

Query Match 61.5%; Score 821; DB 11; Length 468;
Best Local Similarity 37.2%; Pred. No. 1.2e-25;
Matches 154; Conservative 33; Mismatches 58; Indels 169; Gaps 4;

QY 2 WLK-----AFYD----- 4
Db 55 WVKQRPQGLEWIGWIDPDEGETKYAPFQDKATITADTSSNTAYLQLSLTSEDYIYY 114
QY 5 -----AFYD----- 8
Db 115 CARNLLYGGYDYWGQGTITVSSAKTTAPSVVPLAPVCGDTTGSSTVLGCLVKGYFPEP 174
QY 9 -----KVAEKLKAFMDK- 21
Db 175 VTLTNWSSLSGSLGVTFFPAVLQSDLYTLSSSVTVTSSTWPSQSITCNVAHPASSTKVDKK 234
QY 22 -----THTCPP--CAPPELLGGPSVFLFPKPKDITLMISRTPEVTCVVVDVSHEDPEV 72
```

```

235 IEPRGPTIRPCPKCPAPNLLGGPSVFIPPKIKDVLMSLSPMTCVVVDVSEDDPDV 294
QY 73 KFNWYDGVGVHNAKTKPREEQNSTYRVSVLTVLHODWLNKGEYKCKVSNKALPAPIE 132
DB 295 QISWFVNNVEVLTQAQTQTHREDYNSTLRVVSALPIQHODWMSGKEFKCKVNNKALPAPIE 354
QY 133 KTISKAKGQPREQVYTLPPSRDELTKNOVSLTCLVKGFPYSDIAVWESNGQPNENYKT 192
DB 355 RTISKPGSVRAPOVYVLLPPEEMTKQVTLTCMTDWPEDYVETWNNKGTENLYKN 414
QY 193 TTPVLDSGSGFFLYSLKTLVDKSMQOGNVFSCVMHEALHNHYTKSLSPGK 246
DB 415 TEPVLDSGSGFYMSKLRVKEKNWERNYSVSVVHEGLNHHHTKFSRTPCK 468

RESULT 11
Q8R3H6 PRELIMINARY; PRT; 474 AA.
AC Q8R3H6;
DT 01-JUN-2002 (Tremblrel. 21, Created)
DT 01-JUN-2002 (Tremblrel. 21, Last sequence update)
DE Hypothetical 51.7 kDa protein.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RA Strausberg R.;
RL Submitted (MAR-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC025447; AAH25447.1;
KW Hypothetical protein.
SQ SEQUENCE 474 AA; 51748 MW; 8608B57C6DC2874A CRC64;
```

```

Query Match 59.7%; Score 797; DB 11; Length 474;
Best Local Similarity 35.2%; Pred. No. 1.4e-24;
Matches 148; Conservative 40; Mismatches 58; Indels 175; Gaps 4;

QY 1 DWLKA-----FYD----- 5
DB 54 NWKRRPGKLEWIGRIFPGDGTTHYSGKFGQKAKLTADKSSVTAFLQLTSLTSDSAVY 113
QY 6 -----FYD----- 8
DB 114 FCARDSYDGYFDDMGQAGATVTVSSAKTTPPSVYPLAPCGDGTGSSVTLGLVKGYPPE 173
QY 9 -----KVAEKLKEAFMDK 21
DB 174 SVTVTNWSSGLSSVHTFPALLQSLGYTMSSSVTPSPSTWPSQVTCVSAHPASSTTVDK 233
QY 22 -----THTCPP-----CPAELLGGPSVFLPFPKPKDTLMISRTPEVTCVVVDV 65
DB 234 KLEPSGPISTINPCPCCKECHKCPAPNLEGGSVFIFPNKIDVLMISLTPKVTCCVVVDV 293
QY 66 SHEDPEVFNWYDGVGVHNAKTKPREEQNSTYRVSVLTVLHODWLNKGEYKCKVSNK 125
DB 294 SEDDPDVQISWFVNNVEVLTQAQTQTHREDYNSTLRVVSALPIQHODWMSGKEFKCKVNNK 353
QY 126 ALPAPIETISKAKGQPREQVYTLPPSRDELTKNOVSLTCLVKGFPYSDIAVWESNGQ 185
DB 354 DLPSPFIETISKIGLVRAPQVYILPPPAEQLSRKDVSLTCLVGFNPGDISVETNSNGH 413
QY 186 PENNYKTTTPVLDSGSGFFLYSLKTLVDKSMQOGNVFSCVMHEALHNHYTKSLSPG 245
DB 414 TEENKADTAPVLDSGSGFYIYKLDIKTSKWEKTSFSCNVNHEGLKNYLLKTLISRSPG 473
QY 246 K 246
DB 474 K 474
```

```

RESULT 12
Q91Z05 PRELIMINARY; PRT; 473 AA.
AC Q91Z05;
DT 01-DEC-2001 (Tremblrel. 19, Created)
DT 01-DEC-2001 (Tremblrel. 19, Last sequence update)
DE Hypothetical 51.9 kDa protein.
GN AU044919.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RA Strausberg R.;
RL Submitted (JUL-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC010327; AAH10327.1;
DR MGI; MGI:2144967; AU044919.
DR InterPro; IPR003345; CytC_heme_bind.
DR InterPro; IPR003006; Ig_MHC.
DR Pfam; PF00047; Ig_3.
DR PROSITE; PS00190; CYTOCHROME_C; UNKNOWN_1.
DR PROSITE; PS00290; IG_MHC; UNKNOWN_1.
KW Hypothetical protein.
SQ SEQUENCE 473 AA; 51946 MW; CF625F008932AF12 CRC64;

Query Match 59.5%; Score 794; DB 11; Length 473;
Best Local Similarity 41.8%; Pred. No. 1.8e-24;
Matches 148; Conservative 39; Mismatches 58; Indels 109; Gaps 4;

QY 2 WL-----KAFYDK----- 9
DB 120 WLRRIDYWGQGTITVSSAKTTPPSVYPLAPCGDGTGSSVTLGLVKGYPPEVTVTN 179
QY 10 -----VAEKLKEAFMDK----- 21
DB 180 SGLSSSVHTFPALLQSLGYTMSSSVTPSPSTWPSQVTCVSAHPASSTTVDKLEPSGP 239
QY 22 ---THTCPP-----CPAELLGGPSVFLPFPKPKDTLMISRTPEVTCVVVDVSHEDPEV 72
DB 240 ISTINPCPCCKECHKCPAPNLEGGSVFIFPNKIDVLMISLTPKVTCCVVVDVSEDDPDV 299
QY 73 KFNWYDGVGVHNAKTKPREEQNSTYRVSVLTVLHODWLNKGEYKCKVSNKALPAPIE 132
DB 300 QISWFVNNVEVLTQAQTQTHREDYNSTLRVVSALPIQHODWMSGKEFKCKVNNKALPAPIE 359
QY 133 KTISKAKGQPREQVYTLPPSRDELTKNOVSLTCLVKGFPYSDIAVWESNGQPNENYKT 192
DB 360 RTISKIGLVRAPQVYILPPPAEQLSRKDVSLTCLVGFNPGDISVETNSNGTEENYKD 419
QY 193 TTPVLDSGSGFFLYSLKTLVDKSMQOGNVFSCVMHEALHNHYTKSLSPGK 246
DB 420 TAPVLDSGSGFYIYKLDIKTSKWEKTSFSCNVNHEGLKNYLLKTLISRSPGK 473

RESULT 13
Q8WZ42 PRELIMINARY; PRT; 34350 AA.
AC Q8WZ42;
DT 01-MAR-2002 (Tremblrel. 20, Created)
DT 01-MAR-2002 (Tremblrel. 20, Last sequence update)
DE Hypothetical 51.9 kDa protein.
GN TTN.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Freiburg A.; Trombitas K.; Hell W.; Cazorla O.; Fougereousse F.,
```

RA Centner T., Kolmerer B., Witt C., Beckmann J.S., Gregorio C.C.,  
RA Granzier H., Labeit S.;  
RT "Series of exon-skipping events in the elastic spring region of titin  
RL as the structural basis for myofibrillar elastic diversity.";  
RN Circ. Res. 86:1114-1121(2000).  
RP [2]  
RX SEQUENCE FROM N.A.  
RX MEDLINE=21573839; PubMed=11717165;  
RA Bang M.L., Centner T., Fornoff F., Geach A.J., Gotthardt M.,  
RA McNabb M., Witt C.C., Labeit D., Gregorio C.C., Granzier H.,  
RA Labeit S.;  
RT "The complete gene sequence of titin, expression of an unusual ~700  
RT kDa titin isoform and its interaction with obscurin identify a novel  
RT z-line to I-band linking system";  
RL Circ. Res. 89:1065-1072(2001).  
DR EMBL: AJ277892; CAD12456.1; .  
DR InterPro: IPR000282; Cytok\_receptor\_2.  
DR InterPro: IPR000719; Euk\_kinase.  
DR InterPro: IPR000577; FGGY\_kin.  
DR InterPro: IPR003961; FN\_III.  
DR InterPro: IPR001092; HLH\_basic.  
DR InterPro: IPR003599; Ig\_c2.  
DR InterPro: IPR003598; Ig\_c2.  
DR InterPro: IPR003006; Ig\_MHC.  
DR InterPro: IPR003596; Ig\_v.  
DR InterPro: IPR002016; Peroxidase.  
DR InterPro: IPR004168; PPAK\_motif.  
DR InterPro: IPR002290; Ser\_thr\_pkinase.  
DR InterPro: IPR001245; Tyr\_pkinase.  
DR Pfam: PF00041; fn3; 132.  
DR Pfam: PF00047; Ig; 146.  
DR Pfam: PF00069; pkinase; 1.  
DR Pfam: PF02818; PPAK; 53.  
DR ProDom: PD000001; Euk\_pkinase; 1.  
DR SMART: SM00060; FN3; 133.  
DR SMART: SM00409; IG; 167.  
DR SMART: SM00408; IGG2; 148.  
DR SMART: SM00406; IGV; 23.  
DR SMART: SM00220; S\_TKC; 1.  
DR SMART: SM00219; TYR\_KG; 1.  
DR PROSITE: PS00933; FGGY\_KINASES\_1; UNKNOWN\_1.  
DR PROSITE: PS00038; HELIX\_LOOP\_HELIX; UNKNOWN\_1.  
DR PROSITE: PS00290; IG\_MHC; UNKNOWN\_1.  
DR PROSITE: PS00435; PEROXIDASE\_1; UNKNOWN\_1.  
DR PROSITE: PS00111; PROTEIN\_KINASE\_DOM; 1.  
DR PROSITE: PS00109; PROTEIN\_KINASE\_TYR; UNKNOWN\_1.  
SQ SEQUENCE 34350 AA; 3816262 MW; 5B1120058A7CE58A CRC64;

Query Match 36.4%; Score 486; DB 4; Length 34350;  
Best Local Similarity 0.5%; Pred. No. 5,1e-05;  
Matches 164; Conservative 37; Mismatches 44; Indels 30646; Gaps 47;  
Qy 2 WLK----- 4  
|||  
Db 3001 WLKNGVEIKTDKQMRTKLTHSLNIRNVHFGDAADYTFVAGKATSTATLYVEARHIEF 3060  
Qy 5----- 4  
Db 3061 RKHKIDKVLKRAMFECEVSEPDITVQWMKDDQELQITDRIKIQEKYVHRLIPSTR 3120  
Qy 5----- 4  
Db 3121 MSDAGKYTVAGGNVSTAKLFEVGRDVRIRSIKKEVQVIEKQRAVVEFEVNEDDVDAHWY 3180  
Qy 5----- 4  
Db 3181 KDGIEINFQVQERHKYVVERRIRHMFISETRQSDAGEYTFVAGNRSSVTLYVNAPEPPQ 3240  
Qy 5----- 4  
Db 3241 VLQELQPVTVOSGKPARFCAVIGSRPQPKISWYKEEQLLSTGTGFKCKFLHDGQETLLIE 3300  
Qy 5----- 4

Db 3301 AFPEDAAVYTCEAKNDYGVATTSSASLSVEVPEVSWPDQEMPVYPPAIITPLQDVTSEGG 3360  
Qy 5----- 4  
Db 3361 PARFQCRVSGTDLKVSWSYKDKKIKPSRFRMTOPEDTYQLEIAEAYPEDEGTYTFVASN 3420  
Qy 5----- 4  
Db 3421 AVGOVSSPANLSLEAPESILHERIEQETEMEMKEFSSFLSAEERGLHSAEQLQSKINET 3480  
Qy 5----- 4  
Db 3481 LELLSESPYSTKFDSEKGTGPIFKEVSNADISMGDVATLSVTVIGIPKPIQWFFNG 3540  
Qy 5----- 4  
Db 3541 VLLTPSADYKVFVGDGDDHSLIILFTKLEDEGEYTCMASNDYKGTICSAYLKINSKGEGHK 3600  
Qy 5----- 4  
Db 3601 DTETESAVAKSLEKLGPCPPHFLKELKPIRCAOGLPAIFETVVVGEPAITVWFENKQ 3660  
Qy 5----- 4  
Db 3661 LCTSVYVTIIHPNCGSGTFIVNDPOREDSGLYICKAENMLGESTCAEALLVLEDDTMD 3720  
Qy 5----- 4  
Db 3721 TPKAKSTPEAPDFPQTPPLKGPAAVEALDSEQIATFVKDTILKAALITEENQQLSYEHI 3780  
Qy 5----- 4  
Db 3781 AKANELSSQLPLGAQELQSLIQLDQKLTPESTREFLCINGSIHQPLKEPSNQLQIVQS 3840  
Qy 5----- 4  
Db 3841 QKTFSEKIGILPEPETOAVLSDTEKIFPSAMSIQINSLTVEPLKTLAEPEGNYPOSS 3900  
Qy 5----- 4  
Db 3901 IEPPMHSYLTSVAEEVLSPKKTVDNREQRTVLQKQOASALISQSALAEHVESLQS 3960  
Qy 5----- 4  
Db 3961 PDVMSQVNYEPLVPSEHSCTEGGKILLIESANPLENACQDSAVRIEKGSLRFPPLALEEK 4020  
Qy 5----- 4  
Db 4021 QVLLKEHSDNVWMPDQIIESKREPVAIKKQVOYGRDLLSKESLLSGIPPEQRNLKI 4080  
Qy 5----- 4  
Db 4081 QICRALQAASAEQPLFSEWLNRNIEKVEAVNTOEPRHIMCMYLTSAKSVTEVTI 4140  
Qy 5----- 4  
Db 4141 IIEDVDPMANLKMELRDALCAIIEEDILTAEGPRIQOAGKATLQEMDSFSGSQKVE 4200  
Qy 5----- 4  
Db 4201 PITEPEVESKYLISPEEVSYFNVSQSVKYLDPATPVTKGVAASVDSKQESLKPSEKE 4260  
Qy 5----- 4  
Db 4261 ESSSESGTEVATVKIQEAGGFIKEDGPMIHTPLVDTVSEGDIVHLTTITNAKEYNW 4320  
Qy 5----- 4  
Db 4321 YFENKLVPSDEKFKCLQDQNTYTLVIDKVNTEHDGQEVCEALNDSGKATSAKLTVVVKR 4380  
Qy 5----- 4

Db 4381 AAPVIRKKBIEPLEVALGHLAKFTCEIQSAPNVRQFQWKAGREIYESDKCSIRSKYISSL 4440  
QY 5 ----- 4  
Db 4441 EILRTQVDCGEYTCRASNEYGSVCTATITVTEAYPPFLSRPKSLTTFVGKAAFICT 4500  
QY 5 ----- 4  
Db 4501 VTGTPVETIWOQDGAALSPNWKISDAENKHILELSNLTIQDRGVYSKASNKFGADI 4560  
QY 5 -----A 5  
Db 4561 CQBELIIDKPHFIKELEPVQSAINKVKVHLECOVDEDRKVTVTWSDKGOKLPPGDKYKIC 4620  
QY 6 FYDKVA----- 11  
| | | |  
Db 4621 FEDKIATLEIPLAKLDSGTIVCTASNEAGSSCSATVTVREPPSFVKVDPYSLMLPGE 4680  
QY 12 ----- 11  
Db 4681 SARLHCKLKGSPVIQVTFWFKNNKSELSTVRFYNSEAILDITDVKVEDSGYSCEAV 4740  
QY 12 ----- 11  
Db 4741 NDVGSDSCSTEIVIKPEPPFIKLEPADIVRGTNALLQCEVSGTGPFISWFKDKKOIRS 4800  
QY 12 ----- 11  
Db 4801 SKYRLFSQKSLVCLFSPNSADVGEYECVWANEVKGCGCMATHLLKBPFTFVKVDDOL 4860  
QY 12 ----- 11  
Db 4861 IALGGOTVTLQAARVGSEPTVWMKGQEVIREDKIKMSFNGVAVLIIPDVQISFGGK 4920  
QY 12 ----- 11  
Db 4921 YTCIAENAGSQTSGELIVKEPAKIIERAEIQTAGDPATLEYTVAGTPPELKPWKYKD 4980  
QY 12 ----- 11  
Db 4981 GRPLVASKKYRISFKNNVAQLKFYSAEHLDSGQYTFEISNEVGSSCETFTFVLDRDIAP 5040  
QY 12 ----- 11  
Db 5041 FFTKPLRNVDVVNGTCRLDCKIAGSLPMRVSMFKDGEIAASDRYIAFVEGTASLEII 5100  
QY 12 ----- 11  
Db 5101 RVDMDAGNFTCRATNSVGSKDSSGALIVOEPPSFVTKPGSKDVLPGSAVCLKSTFGST 5160  
QY 12 ----- 11  
Db 5161 PLTIRWFGKNKELVSGSCYITKEALESLELYLVKTSBGTYTCVKNVAGGVECSANL 5220  
QY 12 ----- 11  
Db 5221 FVKEPATFVEKLEPSOLLAKGATQACKVTGTPPIKITWANDREIKESKHRMSFVES 5280  
QY 12 ----- 11  
Db 5281 TAVLRLTDVGIEDSGEYMC EAONEAGSDHCSSIVIVKESPYFTKEPKPIEVLKEDVMILL 5340  
QY 12 ----- 11  
Db 5341 AEVAGTPPFEITWFKDNTILRSRGRKKTFIQDHLVLSQLILKFVAADAGEYQCRVTNEVS 5400  
QY 12 ----- 11  
Db 5401 SICSARVTLREPPSFIIKTIESTSLRGTAFAQATLKGSLPITVTWTKDSDEITEDDNIR 5460  
QY 12 ----- 11  
Db 5461 MTFENNVASLYLSGIEVKHDKGVYCOAKNDAGIQRCSALLSVKEPATITEEAVSIDVTG 5520

QY 12 ----- 11  
Db 5521 DPATLOVKFSGTKEITAKWFKDGOELTLGSKYKISVTDIVSILKIISTEKKDKDGEYTFEV 5580  
QY 12 ----- 11  
Db 5581 QNDVGRSCKARINVLDLIIPPSFTKKLAKMDSIKSGFIDLEICIVAGSHPISIQWFKDDQ 5640  
QY 12 -----THTC-----PP----- 27  
| | | | |  
Db 5641 EISASEKYKFSFHDNTAFLEISOLEGTDSGTYTCSATNKAGHNQCSGHLTVKBPYPFVEK 5700  
QY 28 ----- 27  
Db 5701 PQSQDYNPNTRVOLKALVGTAPMTIKWFKDNKELHSGAARSVWKDDTSTLSLELFAKAT 5760  
QY 28 -----C-----PAP----- 31  
| | | | |  
Db 5761 DSGTYICQLSNDVGTATSKATLFVKBPQFIKKPSPVLVRNGQSTTFECQITGTPKIRV 5820  
QY 32 ----- 31  
Db 5821 SWYLDGNEITATQKHGISFIDGLATFOISGARVENSGTYVCEARNDAGTASCISIELKVKE 5880  
QY 32 ----- 31  
Db 5881 PPTFIRELKPVEVKYSDVELECEVTGTPPFEVTLKNNREIRSSKKYTLTDRVSFVNLH 5940  
QY 32 ----- 31  
Db 5941 ITKCDPSDTGEXQCIIVSNEGSCSTRVALKEPPSFIKKIENTTTVLKSSATFQSTVAG 6000  
QY 32 ----- 31  
Db 6001 SPPISITWLKDDQILDDEDDNVYISFVDSVATLQIRSDVNGHSGRYTCQAKNESGVRCYA 6060  
QY 32 ----- 31  
Db 6061 FLLVQEPQAIIVEKAKSDVTEKDPMTLECVAGTPELKVWLKDGQIVPSRVSFMSFEN 6120  
QY 32 ----- 31  
Db 6121 NVASFRIQSVMKDQSGYTFKVENDFGSSCDAYLRVLDQNPSPFTKLTMKDKVLGSS 6180  
QY 32 ----- 31  
Db 6181 IHWEKCVSGLPISAQWFKDGKEISTSAKYRLVCHERSVSLEVNNLEEDTANYTCKVSN 6240  
QY 32 -----ELLGG 36  
| | | | |  
Db 6241 VAGDDACSGILTVKEPPSFLVKPGRQQAIPDSTVEFKAILKGTPTPKIKWFKDDVELVSG 6300  
QY 37 PSVFL----- 41  
| | | | |  
Db 6301 PKCFIGLEGSTSFNLVSDASKTGQYCHVTNDVSDSCSTTMLLVTEPPFKVKLEASK 6360  
QY 42 ----- 41  
Db 6361 IVKAGDSSRLECKIAGSPEIRVWFRNEHELPAOKYRMTFIDSVAVIQMNNLSTEDSGD 6420  
QY 42 -----PPP----- 44  
| | | | |  
Db 6421 FICEAONPAGTSCSTKVIVKEPPVFSPPPIVETLKNAEVSLECELSGTPPEVVMYKD 6480  
QY 45 ----- 44  
Db 6481 KRQLRSSKKYKIAKNFHTSIHILNVDTSDIGEYHCKAQNEVSDTCVCTVKLKEPPREV 6540  
QY 45 -----KPKO----- 48  
| | | | |  
Db 6541 SKLNSLTVVAGEPAELQASIEGAQPIFVQWLKEKEBEVIRESENIRITTFVENVATLQFAKA 6600

QY 49 -----TLM----- 51  
Db 6601 EPANAGYICQIKNDGMEENMATLWLEPAVIVERAGPMVTVGETCTLECKVAGTPEL 6660  
QY 52 ----- 51  
Db 6661 SVEYKDGKLLTSQKHKSFYFNKISSRLILSVERODAGTYTTFQVQNVNGKSSCTAVDV 6720  
QY 52 ----- 51  
Db 6721 SDRAPPSFTRRLKNTGGVILGASCIIECKVAGSPISVAMFHEKTKIVSGAKYQTFSDN 6780  
QY 52 ----- 51  
Db 6781 VCTLQNLSDSDMGNYTCVAANVAGSDECEAVLTVQEPSPFVKPEPEPLEVLPGKNVFT 6840  
QY 52 -----ISRTPEVTCW----- 62  
Db 6841 SVIRGTPPFKVNWFRGARELVKGDRCNIYPEDTVAELEFNIDISQSGEYTCVVSNNAQ 6900  
QY 63 ----- 62  
Db 6901 ASCTTRLFVKEPAFLKRLSDHSVEPGKSIILESTYGTLPISVTWKKGDFNITSEKN 6960  
QY 63 ----- 62  
Db 6961 IVTEKTCILEILNSTKRDAQYSCETIENAGRDVCCALVSTLEPPYFVTEPLEAANG 7020  
QY 63 ----- 62  
Db 7021 DSVSLOQVAGTPEITVSWYKGTKLKRPTEYRTYFTNNVATLVFNKVNINDSGEYTKA 7080  
QY 63 ----- 62  
Db 7081 ENSIGTASSTVFRIORQLPPSFAROLKDIEQTVGLPVLTCRLNGSAPIQVCWYRDGV 7140  
QY 63 -----VDVSHE----- 68  
Db 7141 LLRDDENLOTFSVDNVATLAILQTLDSHGQYSCSASNPLGTASSARLTAREPKKSPFF 7200  
QY 69 ----- 68  
Db 7201 DIRPVSIDVIAGESADFECHVTGAQPMRTWSKDNKEIRPGNVTITCVGNTPHLRILKV 7260  
QY 69 -----DPE 71  
Db 7261 KGDSGOYTQATNDVGKMCQAQLSVKEPPKFKVKKLEASKVAKQGESIQLECKISGPE 7320  
QY 72 VK----- 73  
Db 7321 IKVSWFRNDSSELHESWKYNMSFINSVALLTINEASAFSDSGDYICEAHNGVGDASCPTALT 7380  
QY 74 ----- 73  
Db 7381 VKAPPVFTQKSPVGVKLGSDVILQCEISGTPPPEVWVKDRQVRNKKFKITSKHFT 7440  
QY 74 ----- 73  
Db 7441 SLHLINLEASDVGEYHCKATNEVSGDTCSCSVKFEPPRFVKLLSDTSLIGDAVELRAI 7500  
QY 74 ----- 73  
Db 7501 VEGFQPISVVWLKDRGEVIRESENTRISFDINATLQGSPEASNSGKYICQIKNDAGMR 7560  
QY 74 ----- 73  
Db 7561 ECSAVLTVLEPARIIEKPEMTVTGTGNPFALCECVVTGTPELSAKWFKDGRELSADSKHHI 7620  
QY 74 ----- 73  
Db 7621 TFINVASLKI PCAEMSDKGLYSFEVKNVSGKSNCTVSVHVSRIVPPSFIKRLKDVNAI 7680  
QY 74 ----- 73

Db 7681 LGASVVLECRVSGSAPISVGVFODGNEIVSGPKQSSFSSENVCTNLNLSLLEPSDTGIYTC 7740  
QY 74 ----- 73  
Db 7741 VAANVAGSDECSAVLTVQEPSPFEQTPDSVEVLPGMSLTFTSVIRGTPPFKVKWFKGSR 7800  
QY 74 ----- 73  
Db 7801 LVPGESCNISLEDFVTELELFEVQPLESGDYSCLVTVNDAGSASCTTHLFVKEPATFVKRL 7860  
QY 74 ----- 73  
Db 7861 ADFSVEGTGSPIVLEATYGTGPPISVSWIKDEYLIQSQCSCSITTEKSTILEILESTIED 7920  
QY 74 ----- 73  
Db 7921 YAOVSCLIEAGODICEALVSVLEPPYFIEPLEHVEAVIGEPATLQCKVDGTPEIRISW 7980  
QY 74 ----- 73  
Db 7981 YKEHTKLRSAPAYKMQFKNNVASLVINKVDHSDVGEYSCKADNSGVAGASSAVLVIKERK 8040  
QY 74 -----FNWTVDGV----- 81  
Db 8041 LPPFFARKLKDVBHETLGFPPVAFECRINGSPELVQSVWYKDGVLKKDDANLQTSFVHNVATL 8100  
QY 82 ----- 81  
Db 8101 QILQDQSHIGQXNCASNPLGTASSAKLILSEHVEPPFFDLKPVSDVLDALGESGTFKC 8160  
QY 82 ----- 81  
Db 8161 HVTGTAPKITWAKDNREIRPGGNKMTLVENTATLTVLVKVGKDAGQYTCYASNIAGKD 8220  
QY 82 ----- 81  
Db 8221 SCSAHLGVQEPPEIFIKKLEPSRIVKQDEFTRYECKIGSGPEIKVLWYKDETEIQESSKFR 8280  
QY 82 -----EVHNA----- 86  
Db 8281 MSFVDSVAVLEMHNLNLSVEDSGDYTCEAHNAAGSASSSTSLKVKPEPIFRKKPHIETLKG 8340  
QY 87 ----- 86  
Db 8341 ADVHLECELOGTPPFPHVSWYTKDKRELRSKKYKIMSENFLTSIHILNVDAADIGEYOCKA 8400  
QY 87 -----KTKPR----- 91  
Db 8401 TNDVGSOTCVGSIATKAPPRFVKKLSDISTVVGKEVQLOTTIEGAEPISVWVKDKGEIV 8460  
QY 92 ----- 91  
Db 8461 RESDNIWISYENIATLQFSRVEPANAGYTCQIKNDAGMQECFATLSVLEPATIVBEKPE 8520  
QY 92 ----- 91  
Db 8521 SIKVTTGDTCTLECTVAGTPELSTKWFKDKELTSDNKKYKISFFNKVSGLKIINVAFSDS 8580  
QY 92 ----- 91  
Db 8581 GYVSFEVQNPVGKDSCTASLQVSDRTVPPSFTRKRLKETNGLSGSSVWVECKVYGPPISV 8640  
QY 92 -----EE----- 93  
Db 8641 SWFHEGNEISSGRKYQTLTDNTCALTAVNMLEESDSDGYTCIATNMAGSDECSAPLTVRE 8700  
QY 94 ----- 93  
Db 8701 PPSFVQKPDMDVLVTGNTVFTTSIVKGTPPFSWFKGSSSELVPGDCRNVSLSDSVAELE 8760  
QY 94 ----- 93

Db 8761 LFDVDTQSGETCIYVNEAGKASCTTHLYIKAPAKFVRLNDYSIEKGPLILEGTFGTG 8820  
QY 94 -----QYN----- 96  
Db 8821 TPIPSVTWKNGINVTSPQRNITITTEKSAILEIPSSSTVEDAGQYNCYIENASGRDSCSA 8880  
QY 97 -----STYRV----- 101  
Db 8881 QILILEPPYFVKOLEPVKVSVDGSASLOCLAGTPEIGVSWYKGTDLRPTTTYKMHFRN 8940  
QY 102 ----- 101  
Db 8941 NVATLVFNQVDINDSCEYICKAENSVEVSASTFLTVQEQKLPPSPSROLRDVQETVGLP 9000  
QY 102 -----VS----- 104  
Db 9001 VVFDCAISGEPIVSWSYKDGKPLKDSNNVQTSFLDNTATLNIKFTDRSLAGQYSCATN 9060  
QY 105 ----- 104  
Db 9061 PIGSASSARLILTEGKNPPFFDIRLAPVDVAVGESADFECHVTGTQPIKVSNAKDSREI 9120  
QY 105 -----LTVLHOD-----WLNKGE----- 117  
Db 9121 RSGGKQIISYLENSAHLTVLKVDKGDGSGQYTCYAVNEVGKDSCTAQLNIKERLIPPSTFK 9180  
QY 118 ----- 117  
Db 9181 RLSETVEETEGNSFKLEGRVAGSQPITVAVYKNNIEIQTSNCEITFKNNTLVQVRKAG 9240  
QY 118 -----YKCKVSN----- 124  
Db 9241 MNDAGLYTCKVNDAGSALCTSSIVIKPKPPVDFDHLTPVTSEGEYVQLSCHVQSGE 9300  
QY 125 ----- 124  
Db 9301 PIRIOWLKAGREIKPSDRCSFSFASCTAVLELRDVAKADSGDYVCKASNVAGSDTTKSKV 9360  
QY 125 ----- 124  
Db 9361 TIKDRPAVAPATKAAVDGRLEFFVSEPOSIRVVEKTATFIKVGGDPIPNVKTGKWR 9420  
QY 125 ----- 124  
Db 9421 QLNQGRVFIHQGDGALEIRDTTDSGLYRCVAFNEHGEIESNVNLOVDERKKQEKI 9480  
QY 125 ----- 124  
Db 9481 EGDRLMLKKTPIILKKGAGEEEDIMELLKNVDPKVEKYARMYGITDFRGLLOAFELL 9540  
QY 125 ----- 124  
Db 9541 KQSQEETHRLIEBIERSERDEKEFEELVFSIQORLSQTEPVTILKDIENTQVLKNDNA 9600  
QY 125 ----- 124  
Db 9601 VFEIDIKINYPEIKLSWYKTEKLEPSDKFEISIDGRHRTLKVKNCQLKDQGNRYLVCGP 9660  
QY 125 ----- 124  
Db 9661 HIASAKLTVIEPAWHERHLODVTLKEGQCTWTCQFSVPNVKSEWPRNGRIILKPOGRHTE 9720  
QY 125 ----- 124  
Db 9721 VEHKVHLTIADVRAEDQGYTCKYEDLETSAELRIEAEPIQFTKRIONIVVSEHQSATF 9780  
QY 125 ----- 124  
Db 9781 ECEVSFDDAIVTWYKGTLETESQYNFRNDGRCHYMTIHNVTPDDEGVSVIARLEPRG 9840  
QY 125 ----- 124  
Db 9841 EARSTAELYLTKEIKLELKPDPIDPSRVPIPTMPIRAVPEEIPPPVAPPILPLLTPE 9900

QY 125 ----- 124  
Db 9901 EKKPPPRIEVTKAVKDKDAKKVAKPKEMTPREEIVKKPPPTTLIPAKAPEIIDVSSK 9960  
QY 125 ----- 124  
Db 9961 AEEVKIMTITRKKEVQKEAVEYKQAVHKRVFIESFEPYDELEVETPEPQY 10020  
QY 125 ----- 124  
Db 10021 YEEPDEDEYEEKVAKEVHEEWEEDFERGQYERYEREGYDEGEEREAYOEREVIQVO 10080  
QY 125 ----- 124  
Db 10081 KEVYEESHERKVPKAVEKKAPPPKVIKKPVIEKIEKTSRRMEBEKVQTVKVPESKKI 10140  
QY 125 ----- 124  
Db 10141 VPQKPSRTPVQEEVIEVKVPAVHTKKMVISEKMFASHTEESESVTVPEVQKEIVTEBK 10200  
QY 125 ----- 124  
Db 10201 IHVAVSKRVPPKVPPELPEKPAPEEAPVPIPKKVEPPAPKVPPEPKKVPPEKKVPV 10260  
QY 125 ----- 124  
Db 10261 PKKEPAAPPKVPPEVKPPVEEKIPVPVAKKKEAPPKVPVEQKRVVTEEKITIVTQREE 10320  
QY 125 ----- 124  
Db 10321 SPPPAVPEIPKKVPPEERKVPVRKEEVPVPKVPALPKVPPEKVAVVPVPAKKAPPP 10380  
QY 125 ----- 124  
Db 10381 RAEVSKKTVEEKRFVAEEKLSFVQPVRVETRHEVSAAEESEWSYSEEEGVSISVYREE 10440  
QY 125 ----- 124  
Db 10441 REEEBAEYVEVMEEPPEEYVVEEKLHII SKRVEAEPAEYTEROEKKIVLKPAPAKIE 10500  
QY 125 -----KALPAPI----- 131  
Db 10501 EPPPAKVPAPKKIVPEKKVPAPVPKKEKVPPEKPPKVPPEKPPKVPKVPKKEEPLP 10560  
QY 132 ----- 131  
Db 10561 AKVTERHMOITQEEKVLVAVTKKEAPPKARVPEEPKRAVPEEKVLKLPKREEPAPKVT 10620  
QY 132 ----- 131  
Db 10621 EFRKRVKKEKVSIEAPKREPOPIKEVTIMEEKERAYTLEBEANSVQREEEYEEYDY 10680  
QY 132 ----- 131  
Db 10681 KEFEYEPTEDQYEEYEREYERYESHEEYITEPEKPIPVKVPPEEPVPTKAPDAK 10740  
QY 132 ----- 131  
Db 10741 VLKKAPEEKVPVPIPKKLKPPPPKVPPEPKVPPEKTRISITTKREKQVTEPAKAVPMK 10800  
QY 132 ----- 131  
Db 10801 PKRVVAEEKVPVRKEVAPPVVRVPEVPKLEPPEEVAFFEEVVTHVEEYLVSEEEYIH 10860  
QY 132 ----- 131  
Db 10861 EEFITEEVPVPIPVKVPPEVRKVPPEKKVPVPKKEAPPKVPPEPKKVPPEKVPVLI 10920  
QY 132 ----- 131  
Db 10921 PKKEKPPPAKVPPEVKPPVPEEKVPVPVPPKVEAPPKVPPEPKKVPPEKVPVAPKVV 10980

Qy 132 ----- 131  
Db 10981 EAPPAKVEVPKLLIPEKKKTPVPKVEAPPKVPKREPVPVVALPQEEVLFEEI 11040  
Qy 132 ----- 131  
Db 11041 VPEEVLPEEEVLPEEEVLPEEEIPEEEIPEEEVPEEEVPEEEV 11100  
Qy 132 ----- 131  
Db 11101 LPEVKVPVPAPVPEIKKTEKKVIPKKEAPPKVPKVEKRIILPKEEVL 11160  
Qy 132 ----- 131  
Db 11161 PVEVTEPEEPISEEEIPEEPPSIEEVEVAPPRVPEVIKAVPEAPTVPKKEAPPA 11220  
Qy 132 ----- 131  
Db 11221 KVSKKIPEKVPVPVORKEAPPKVPKVPKPEKVLVPKKEAVPPAKGRTVLEEKVS 11280  
Qy 132 ----- 131  
Db 11281 AFROEVVVKERLEVEVAEVEEIPPEEPEFHEVEYFEEGFEHEVEEPIKLEQHRVEEH 11340  
Qy 132 ----- 131  
Db 11341 RVEKVRVIEVPEAEVEVFEKPKAPKGPSEIKIIPPKKPPTKVVPKPEPPKAPVPEV 11400  
Qy 132 ----- 131  
Db 11401 KKIIVEEKVRVPEBPRVPPTKVPDVLPPKEVPEKKVPPPAKPEAPPKVPAPKEV 11460  
Qy 132 ----- 131  
Db 11461 PEKKVPPPPKKPEVPTTKVPEVKAAPVEKKVPEAIPPKPEPPPEVPEAPKEVVPPEK 11520  
Qy 132 ----- 131  
Db 11521 VPAAPPKKPEVTPVKVPEAPKEVVPKVPVPPPKKPEVPTTKVPEVPAVPEKKVPEA 11580  
Qy 132 ----- 131  
Db 11581 IPPKPEPPPEVEEVEALEEPAEVEPEPAAPQVTPVPKVPPEKAPAVAKK 11640  
Qy 132 ----- 131  
Db 11641 PELPPVKVPEVPEVPEKKVPLVVPKPEAPPKVPPEVPEVPEKKVAVPPKKPEVPPA 11700  
Qy 132 ----- 131  
Db 11701 KVPEVPPKVPLEKPAVPVPRAESPPPEVPEEPEEIAPEEIEAPEEKPVPVAAEEPE 11760  
Qy 132 ----- EKTISK----- 137  
Db 11761 VPPPAVPEPKIIPKVKVPVVKKPEAPPKPEPEKVIKPKLPKPPPPPPPPPKEDV 11820  
Qy 138 ----- 137  
Db 11821 KEKIFOLKAIKKVKPKVPQVPEKVELTPLKVPGEKKVKRLLPERKPEEVEVLKSVL 11880  
Qy 138 ----- 137  
Db 11881 RKRPEEPEPKVKLEKVKKPAVPEPPPPKPEVEVEVPTTKRERKIPEPTKVPEIKPA 11940  
Qy 138 ----- 137  
Db 11941 IPLPAPEPKPEAEVTKIKPPVPEPEPTIAAPVTPVVGKAEAKAPEAKPKGPI 12000  
Qy 138 ----- 137  
Db 12001 KGVPKTSPTEAERRKLPGSGGKPPDEAPFTYQLKAVPLKFVKEIKDIILTESEFVG 12060  
Qy 138 ----- 137

Db 12061 SSAIFECVLSFSTALTTHMKDGSNIRESPKHRFIADGKDRKLHIIDVOLSDAGEVTCVLR 12120  
Qy 138 ----- AKGP----- 142  
Db 12121 LGNKEKSTAKLVVEELPVRFVKTLEEBVTVMKGPLYLSCELNKRVDVVRKDGKIIVE 12180  
Qy 143 ----- 142  
Db 12181 KPGRIVPGVIGIMRALTINDADDDTAGTYTTVTENANNLECCSVKVVEVIRDWLVKPIR 12240  
Qy 143 ----- 142  
Db 12241 DOHVKPKGTALFACDIADKTPNIKWFKGDEIPAEPNDKTEILRDGNHLYLKIKNAMPED 12300  
Qy 143 ----- 142  
Db 12301 IAEYAVEIEGKRYPAKLTGGERVELLKPIEDVTIYKESAFDAEISEADIPGQWLKG 12360  
Qy 143 ----- 142  
Db 12361 ELLRPSPTCEIKAEKGKRFLTLRKVKLDQAGEVLYQALNAITAILTVKEIELDFAVPLK 12420  
Qy 143 ----- 142  
Db 12421 DVTVPERQARFECVLTREANVINSGPDIKSSDKFDIIADGKKHILVINDSQFDEGV 12480  
Qy 143 ----- 142  
Db 12481 YTAIEVEGKTSARLFTVGTIRLKFMSPLEDOTVKEGETATFVCELSHERKMHVWFKNDAKL 12540  
Qy 143 ----- 142  
Db 12541 HSTRVTLISSEGTHKLEMKVETLDDISQIRAOVKELSSAQKLVLEADPTFTVKLHDKT 12600  
Qy 143 ----- 142  
Db 12601 AVEKDEITLKCEVSKDVPVKWFKGDEIIVSPKYSIKADGLRRILKIKKADLKDKGEVVC 12660  
Qy 143 ----- 142  
Db 12661 DCGTDTKANVTVEARLIKVEKPLYGVFVGETAHTAEIELSEPDVHGQWLKGQPLTAS 12720  
Qy 143 ----- 142  
Db 12721 PDCEIIEDGKKHILLHNCQLGTMGEVSFQANAKSAANLKVKEPLPLIETPLSDVKVFE 12780  
Qy 143 ----- REPO----- 146  
Db 12781 KDEAKFECEVSRPKTFRWLKGTQETITGDRFELIKDGTKHSWIKSAAFEDEAKYMFEA 12840  
Qy 147 ----- 146  
Db 12901 VSMQDEGKTHSITFKDLSIDDTSQIRVEAMGMSSEAKLTVLEGDPTFTGKLQDTGVKED 12960  
Qy 147 ----- 146  
Db 12961 EVILQCEISKADAPVKWFKDGKEIKPKNAVIKADGKKRMLLKKALKSDIGQYTCDCGT 13020  
Qy 147 ----- 146  
Db 13021 DKTSGKLDIEDREIKVRLPHSVEVMETETARFETIESEDITHANWKLGEALLQPDCE 13080  
Qy 147 ----- 146  
Db 13081 IKEEGKIHSVLHNCRLDQOTGVDFQAANVSSAHLRVKPRVIGLLRPLKDVTVTAGETA 13140  
Qy 147 ----- 146

Db 13141 TFCELSYEDIPVEWYLGKKLEPSDKVVRSEKGVHTLFLRDVKLEDAGEVOLTAKDFX 13200  
QY 147 ----- 146  
Db 13201 THANLFVKEPPVETKPLEDOTVEEGATAVLECEVSRENAKVWFKNGTIELKSKKYEIV 13260  
QY 147 ----- 146  
Db 13261 ADGRVRKLVHDCTPEDIKYITCDAKDKFTSCNINVPPHVEFLRPLTDLQVREKEMARF 13320  
QY 147 ----- 146  
Db 13321 ECELSRENAKVKWFKDGAIEIKGKYDIISKGAVRILVINKCLLDDAEAYSCVETARTS 13380  
QY 147 ----- 146  
Db 13381 GMLTVLEEEAVFTKNLANIEVSETDIKLVEVSKPGEAVIWKYGDDEIIETGRYEILTE 13440  
QY 147 ----- 146  
Db 13441 GRKRILVIONAHLEDAGNYCNRLPSSRTDGVKVKVHELAAEFISKPNLEILEGEKAFFVC 13500  
QY 147 ----- 146  
Db 13501 SISKESFPVQMKRDKLTLESGDKYDIADGKKRVLVVKDATLQDMGTYYVMVGAARAAH 13560  
QY 147 ----- 147  
Db 13561 LTVIEKLRVVPLKDRVRKEQEVFENCEVNTGAKAKWFRNEEAIFDSKYIILQKOLV 13620  
QY 148 YTL----- 150  
|||  
Db 13621 YTLRIRDAHLDDQANYNSLTNRGENVKSAANLIVEEEDLRIVEPLKDIEIETMEKKSIVF 13680  
QY 151 ----- 150  
Db 13681 WCKVNRNLNVLKWTNGEEVFPDNRVSYRVDYKHKMLTIKDCGFPDEGEYIVTAGODKSV 13740  
QY 151 ----- 150  
Db 13741 AELLIIIEAPTEFVEHLEDQTVTEFDDAVFSCQLSREKANVKWYRNGREIKGKKYFEKD 13800  
QY 151 ----- 150  
Db 13801 GSIHRLIIKDCRLDDEYACGVEDRSRFLFVEIIPVQDILEAPGADVFLA 13860  
QY 151 ----- 150  
Db 13861 ELANKVEVOWLRNNMVVQGDHKQMMSEGIHRLQICDIKPRDQGEYRFIAKDEKARAK 13920  
QY 151 ----- 150  
Db 13921 LELAAAPKIKTADQDLVDVVGKPLTWVPYDAYPKAAEWFKENEPISKTIDTTAEQTS 13980  
QY 151 ----- 150  
Db 13981 FRILEAKGDKGRYKIVLQNHKGAEFNLKVIDVPGVPRNLEVTEFDEGEVSLAWEPP 14040  
QY 151 ----- 150  
Db 14041 LTDGGSKIIGYVVERRDIKRKTWVLATDRAESCEFTVTGLQGGVEYLFVRSARNRVGTG 14100  
QY 151 ----- 150  
Db 14101 EPVETDNPEARSKYDVPGLPNVTITDYNRFVSLTWEPPEYDGAETINVIYIELRDKT 14160  
QY 151 ----- 150  
Db 14161 SIRWDTAMTVRAEDLSATVTVVEGOEYSFVRQAQRNIGVGKPSAATPEVKVADPIERPS 14220  
QY 151 ----- 150  
Db 14221 PPNVLTSSDOTQSSVOLKWEPPPLKDGSGPILGYIIRCEEGKDNWJRCNKLVPETYKV 14280

QY 151 ----- 150  
Db 14281 TGLEKGNKLYRYSAENKAGVSDPSEILGLPTADDAFVEPTMDLSAFKDCGLEIVPNPIT 14340  
QY 151 ----- 150  
Db 14341 ILVPSTGYPRPTATWCFGDKVLETGDRVKMKTLSAYAEVLSPSERSDKGIYTLKLENRV 14400  
QY 151 ----- 150  
Db 14401 KTISGEIDVNVIAIARPSAPKELKFGDITKDSVHLTWEPDDDGSPLTGYVVEKREVSRT 14460  
QY 151 ----- 150  
Db 14461 WTKVMDFTVDTLETPVDLVQKKEYLFKVCARNKCGGEPAYVDEPVNMSTPATVPDPEN 14520  
QY 151 ----- 150  
Db 14521 VKWRDRTANSIFLTWDPKNDGSGRIKGYIVERCPRGSKWACGEPVAETKMEVTGLEE 14580  
QY 151 ----- 150  
Db 14581 GRWYAYRVKALNRQGASKPSRPTTEIQAVDQEAPEIFLDVKLLAGLTVKAGTKIELPAT 14640  
QY 151 ----- 150  
Db 14641 VTGKPEPKITWTAKADMILKQDKRITIENTVNPVKSTVTIVDSKRSDDTGTYYIIIEAVNVCGRAT 14700  
QY 151 ----- 150  
Db 14701 AVVEVNVLDKPGPAADFITDVTNESCLLTNPPRDDGSKITNYVVERRATDSEVWHKL 14760  
QY 151 ----- 150  
Db 14761 SSTVKDTNFKATKLIPNKEYIFRVAENMYGVGEVQASPIATAKYQDPPGPTRLEPSD 14820  
QY 151 ----- 150  
Db 14821 ITKDAYTLTWCEPDDDDGSGPITGYVVERLDPTDKWRCNKMPVKDITTVRVKGLTNKKY 14880  
QY 151 ----- 150  
Db 14881 RFRVLAENLAGPKPSKSTEPILIKDPIPPWPPGKPTKDVGKTSVRLNWTKEHDGGA 14940  
QY 151 ----- 150  
Db 14941 KTESYVIEMLKTGTDENVRAEGVPTTQHLLPGLMEGQEYSFRVRAVNKAGESEPEPSD 15000  
QY 151 ----- 150  
Db 15001 PVLCREKLYPPSPRRWLEVINITKNTADLKWTVPKDGSGSPITNYIVEKRDVRKCMQTV 15060  
QY 151 ----- 150  
Db 15061 DTTVKDKTCTVPTLEBSLYFVRVAENAIGOSDYTEIEDSVLAKDFTTTPGPPYALAVV 15120  
QY 151 ----- 150  
Db 15121 DVTKRHRVDLKWEPKNDGGRPIQRYVIEKKERLGTNRWVAGKTAGPDCNFRVTDVIEGTE 15180  
QY 151 ----- 150  
Db 15181 VOQVRAENEAGVGHSEPTEILSIEDTSPSPPLDLHVTADGRKHIAIANKPPEKNG 15240  
QY 151 ----- 150  
Db 15241 SPIIGYHVEMCPVGTGKWRVNSRPIKDLKFKVEGVVDPKEYVLRVAVNAIGVSEPS 15300  
QY 151 ----- 150  
Db 15301 ISENVAKDPDCKPTIDLETHDIIIEGKLSIPVFPRAVPVTVSNHKGKVKASDRL 15360



QY 151 ----- 150  
Db 15361 TMKNDHLSAHEVPKSVRADAGIYITITLENKLSATASINVKIGLPGCKDIKASDITK 15420  
QY 151 ----- 150  
Db 15421 SSCKLTWEPPEFDGGTPIHLVLERREAGRRTYIPVMGENKLSWTVKOLIPNGEYFFRV 15480  
QY 151 ----- 150  
Db 15481 KAVNKVGGGEYIELKNPVIADQPKQPPDPVVEVHNPTAEAMTITWKPLYDGGSKIMG 15540  
QY 151 ----- 150  
Db 15541 YIIETIAKEERNKRCNEHLVPLITYTAKGLEBGEKQFVRVRAENAAIGSEPSRAPPTK 15600  
QY 151 ----- 150  
Db 15601 AVDPIDAPKVILRTSLEVKRGDEIALDASISGSPYPTITWIKDENVIVPEIKKRAAPLV 15660  
QY 151 ----- 150  
Db 15661 RRRKEVOEEPPFVLPTQRLSIDNSKKGESQLRVRDSLDPDHGLYMIKVENDHGIAKAP 15720  
QY 151 ----- 150  
Db 15721 CTVSVLDTGPPINFVFEDIRKTSVLCKWEPPLDDGSGSEIINTLEKKDKTKPDSEWIVV 15780  
QY 151 ----- 150  
Db 15781 TSTLRHKYSVTKLIEGKEYLFRVRAENRFGPGPCVKPLVAKDPFGPPDAPDKPIVED 15840  
QY 151 ----- 150  
Db 15841 VTSNMLVKWNEPKDNGSPILGYWLEKREVNSTHWSRVNKSLLNALKANVDGLLEGLTYV 15900  
QY 151 ----- 150  
Db 15901 FRVCAENAGPGKFPSPDKTAHPISPPGPIPRVTDTSSTIELEWEPFANGGEI 15960  
QY 151 ----- 150  
Db 15961 VGYFVDQLVGTNEWSRCTEKMIKRVQYTVKEIREGADYKLRVSANAAEGEPGETQPV 16020  
QY 151 ----- 150  
Db 16021 TVAEPPQPAVELDVSVKGGIOMAGKTLRIPAVVTGRPVPTKVWTKEGELDKDRVID 16080  
QY 151 ----- 150  
Db 16081 NVGTSKELIIKDALRKDHGRVITATNSCGSKFAAARVEFDVPGPVLDLKPVVTRKMC 16140  
QY 151 ----- 150  
Db 16141 LLNWSDEDDGGSEITGTIIERKDAKHTWROPIETERSKCDITGLLEGQYKFRVIK 16200  
QY 151 ----- 150  
Db 16201 KFGGPPVEIGPILAVDPLGPPTSPERLTYTERTKSTITLDWKEPRNGSGSPIQGYIIEK 16260  
QY 151 ----- 150  
Db 16261 RRHDKPDFERNKRLCPTTSFLVENLDEHOMYEFVRKAVNEIGESEPSLPINVVIOQDEV 16320  
QY 151 ----- 150  
Db 16321 PPTIKRLSVRGDTIKVAGEPVHIPADVTGLPMPKIEWSKNETVIEKPTDALQITKEEV 16380  
QY 151 ----- 150  
Db 16381 SRSEAKTELSIPKAVREDKCTYTVTASNRLGVSFRNVHVEYDRPSPRNLAVIDIKAES 16440  
QY 151 ----- 150

Db 16441 CYLTWAPLDNGSGSEITHYVIDKRDASRKKAEWEVINTAVEKRYGIWKLI PNGQVEFRV 16500  
QY 151 ----- 150  
Db 16501 RAYNKYGISDECKSDKVVIQDPYRLPGPGPKVLARTKGSMLVSWTPPLDNGGSPITGY 16560  
QY 151 ----- 150  
Db 16561 WLEKREGSPYWSRVSRAPITKVGLKGVEFNVPRLLEGVKYQFRAMAINAAGIGPPSEPS 16620  
QY 151 ----- 155  
|:::|  
|:::|  
Db 16621 DPEVAGDPIPPGPPSPCEVKDKTKSSISLGWKPPAKDGGSPIKGYIVEMQEETTDWKR 16680  
QY 156 ----- 155  
Db 16681 VNEPDKLITTCCECVPNLRELKRYRFRKAVNEAGESEPSDTTGEIPATDIOEPEVFID 16740  
QY 156 ----- 155  
Db 16741 IGAQDCLVCKAGSQIRIPAVIKGRPTPKSSWEFDGKAKKAMKDVHDIPEDAQLETAENS 16800  
QY 156 ----- 155  
Db 16801 SVIIIPCKRSHTGKYSITAKNAGOKTANCRVKVMDVPKDLKVSIDITRGSCLSWK 16860  
QY 156 ----- 155  
Db 16861 MPDDGGDRIGYVIEKRTIDGKAWTKVNPDCGSTTFVVPDLJSEQQYFFRVAENRFGI 16920  
QY 156 ----- 155  
Db 16921 GPPVETIQTARTDPIYPPDPPIKLKIGLITKNTVHLSWPKPKNDGSGSPVTHYIVECLAW 16980  
QY 156 ----- 155  
Db 16981 DPTGTKEAWROCNKRDRVEELQFTVEDLVEGGEYFRVKAVNAAGVSKSPSATVGPCDCQR 17040  
QY 156 ----- 155  
Db 17041 PDMPPSIDLKEFMEVEBGTNVNIVAKIKGVPPPTLTWFKAPPKPNKEPVLVDYHVNKL 17100  
QY 156 ----- 155  
Db 17101 VDDTCTLVIPQSRSDTGLYTTITAVNNLGTASKEMRLNLVLRPGPPVGPPIKPFESVADQ 17160  
QY 156 ----- 155  
Db 17161 MTLSPFPKDDGSKITNYVIEKREANRKTWVHVSSEPKECTYTI PKLLEGHEVYFRIMA 17220  
QY 156 ----- 155  
Db 17221 QNKYIGEPLDSEPTARNLFSVPGADKPTVSSVTRNSMTVNWEEPEYDGGSPVTGYWL 17280  
QY 156 ----- 155  
Db 17281 EMKDTTSKRWRNRDPIKAMTLCVSYKVTGLIEGSDYQFVRVAINAAGVGPASLPSDPA 17340  
QY 156 ----- 155  
Db 17341 TARDPIAPGPPPKVTDWTWKSSADLEWSPPLKDDGSKVTCYIYVEYKEEGKEEWEKCKDK 17400  
QY 156 ----- 155  
Db 17401 EVRGTKLVVTGLKEGAFYKFRVSAVNIAGICEPGEVTDVIEMKDRLVSPDLQDASVRDR 17460  
QY 156 ----- 155  
Db 17461 IVVHAGVIRIIAYVSGKPPPTVTWNNERTLPQEATIIETAISSSVIKNQORSHGVY 17520  
QY 156 ----- 155

Db 17521 SLLAKNEAGERKKTIIVDVLDPGPVGTPLAHLNLTNESCKLTWFSPEDDGGSPITNYVI 17580  
QY 156 ----- 155  
Db 17581 EKRESORRAWTPYTYVTQRNATVOGLIOCKAYFFRIAAENSGMGPFVETSEALVIREP 17640  
QY 156 -----ELTKNOVSLT----- 165  
Db 17641 ITVPERDELEVKEVTKNTVTLTNWPKYDGGSEIINYVLESRLIGTEKFKVTNDNLLS 17700  
QY 166 -----CLV- 168  
Db 17701 RKYTVKLGKEDGYEYRSVAVNIVGOGKPSFCTKPTCKDELAPPTLHLDKRLTIRVG 17760  
QY 166 -----CLV- 168  
Db 17761 EAFALTGRYSGPKPKVSKFKEADVLEDDRTHIKTTPATLALEKTKAKRSDSGKYCVV 17820  
QY 169 -----KGF----- 171  
Db 17821 ENSTGRKGFQCVNVVDRPGPPVGVSFDEVTKDYMVISWKPPLDGGSKIITNYIEKKE 17880  
QY 172 ----- 171  
Db 17881 VGKDVMPVTSASAKTTCCKVSKLEGKDYIFRIHAENLYGIDPLVSDSMKAKDRFRVPD 17940  
QY 172 ----- 171  
Db 17941 APDOPIVTEVKDSALVTWKNKPHDGKPTITNYILEKRETMKRWARVTKDPIHPTKFRV 18000  
QY 172 ----- 171  
Db 18001 POLLEGQYEFVRSVAENEIGDPSPPKPVFAKDPFAKPPVNPPEAIDTTCNSVDLTW 18060  
QY 172 ----- 171  
Db 18061 QPPRHGGSKILGYIYEVQKVGDEWRANHTPESCPETKYKVTGLRGDQTYKFRVLAVN 18120  
QY 172 ----- 171  
Db 18121 AAGESDPAHVPEPVLVKDRLEPPELLIDANMAREQHIKVGDTLRLSAILIKGVFPFKVTWK 18180  
QY 172 ----- 171  
Db 18181 KEDRAPTARIDVTPVGSKEIRNAAHEDGGIYSLTVENPAGSTVSVKVLVDKPGPP 18240  
QY 172 ----- 171  
Db 18241 RDLVSEIRKDCSYLTWKEPDDGGSVITNYVVERRDVAASQWSPLSATSKKSHFAKHL 18300  
QY 172 ----- 171  
Db 18301 NEGNQYLFRVAENQYGRGPFVETPKPKALDPLHPPGPKDLHHVDVDKTEVSLVWNKP 18360  
QY 172 ----- 171  
Db 18361 DRDGGSPITGYLVEYQEGTDWIKFKVTNLECVVTGLQOGKTYRFRVKAENIVGLGLP 18420  
QY 172 ----- 171  
Db 18421 DTTPIEQEKLVPSPVELDVKLEGLVVVKGACTTVRFPAIIRGVVPVPTAKWTTDGESEIKT 18480  
QY 172 ----- 171  
Db 18481 DEHYTVETDNFSSVLTIKNCLRDTEGYOITVSNAAGSTVAVHLTVLDVPGPTGPINI 18540  
QY 172 ----- 171  
Db 18541 LDVTPHEMTISWQPPKDDGGSPVINYIVEKQDTRKDTWGVSSGSKTKLIPHQLKGCE 18600  
QY 172 ----- 171  
Db 18601 YVFRVRAENKIGVGPPLDSTPTVAKHKFSPSPGKPVVTDITENAAVTSWTLPSKDGGS 18660

QY 172 ----- 171  
Db 18661 PITGYMERREVTGKVRVNKTPIALDKFRVTGLYEGNTYEFVRVFAENLAGLSKPSSTD 18720  
QY 172 ----- 171  
Db 18721 PIKACRPIKPPGPPINPKLKDKSRETADLVWTKPLSDGGSPILGYVVECOKPTCAQWNR 18780  
QY 172 ----- 171  
Db 18781 NKDELITROCAFVRPGLIEGNEYRFRRIKAANIVGEPEPRELAESVIAKIDILHPPEVELDVT 18840  
QY 172 ----- 171  
Db 18841 CRDVTVRVGQITIRILARVKGRPEPDITWTKEGKVLVREKRVLDLIQDLPRVELQIKEAVR 18900  
QY 172 ----- 171  
Db 18901 ADHGKVIISAKNSSGHAQSAIVNVLDPRGPCQNLKVTNVTKENCCTISWENPLDNGSEI 18960  
QY 172 ----- 171  
Db 18961 TNFIVEYRKPNOGWSIVASDVTKRLIKANLLANNEYFRVCAENKVGVTIETKPTIL 19020  
QY 172 ----- 171  
Db 19021 AINPIDRPGEPENLHIADKGKTFVYLKWRPPDYDGGSPNLSTYHVERRLKGSDDWERVHKG 19080  
QY 172 ----- 171  
Db 19081 SIKETHYMYDRCVENQIYEFVRQTKNEGESDWWKTEVVVKEDLQKPVLDLKSGLVTV 19140  
QY 172 ----- 171  
Db 19141 KAGDTIRLEAGVRGKPFPEVAWTKDKDADTLTFRSPRVKIDTRADSKFSLTAKRSDGOK 19200  
QY 172 ----- 171  
Db 19201 YVVATNTAGSFVATVNVLDKPGVNRNLKIVDVSSDRCTVCWDPPEDDGGCEIQNYIL 19260  
QY 172 ----- 171  
Db 19261 EKCETKRMVWSTYSATVLTPTGTTVTRLIEGNEYIFRVAENKIGTGPPTESKPVIATKY 19320  
QY 172 ----- 171  
Db 19321 DKPGRDPPEVTKVSKEEMTVVWNPPEYDGGKSITGYFLEKKEKHSTRVWPVNKSAIPER 19380  
QY 172 ----- 171  
Db 19381 RMKVQNLPLPDHEYQFRVKAENEIGIGEPSLPSRPVVAKDPIEPPGPTNFRVVDTTKHSI 19440  
QY 172 ----- 171  
Db 19441 TLGWKPVYDGGAPIIGYVVENMRPKIADASPOEGWKRCAAAQLVRKEFTVTSLDENOEY 19500  
QY 172 ----- 171  
Db 19501 EFRVCAQNOVGIRPAELKEAIKPKKEILEPPEIDLDASMRKLVIVRAGCPIRLFAIVRGR 19560  
QY 172 ----- 171  
Db 19561 PAKVYTWKRGIDNVVRKQVDLVTMAFLVTPNSTRDDSGKYSLLVNPAGEKAVFVNV 19620  
QY 172 ----- 171  
Db 19621 RVLDTGPPVSLKVSADVTKTSCHVSWAPPENDGGSQVTHYIVEKREADRKTWSTVTPEVK 19680  
QY 172 ----- 171  
Db 19681 KTSFHVNTLVPNGNEYFRVTVAVNEYGPGVPTDVPKPVLASDPLSEDPDRPKLEVTEMKN 19740

QY 172 ----- 171  
Db 1741 SATLAWPLRLDGGAKIDGYITSYREEEQPADRWTEYSVVKDLSLVVTGLKEGKKYFRV 19800  
QY 172 ----- 171  
Db 19801 AARNAVGLPREAGVYEAKEQLLPKILMPEOITIKAGKKURIEAHVYGKPHPTCKWK 19860  
QY 172 ----- 171  
Db 19861 KGEDEVTSHLAVHKADSSILLIKDVTBKDSGYISLTAENSSGTDQIKVVMMDAPG 19920  
QY 172 ----- 171  
Db 19921 PPQPPFIDISDADACSLSWHIPLEDDGGSNITNIVEKCDVSRGDWVTALASVTKTSRV 19980  
QY 172 ----- 171  
Db 19981 GKLIPGOEYIFRVAENRFGISEPLTSPKMAOFFPGVPSEPKNARVTKNKDCIFAWD 20040  
QY 172 ----- 171  
Db 20041 RPDSDGSGPIIGYLIERKERNLLWVKANDTLVRSYPCAGLVEGLEYSFRIYALNKAG 20100  
QY 172 ----- 171  
Db 20101 SSPSPKTEYVARMVPDPPCKPEVIDVTKSTVSLIWARPKHDGSGKIIIGVFEACKLPG 20160  
QY 172 ----- 171  
Db 20161 DKWVRTAPHQIPQEEYTATGLEEKAQYQFRAIARTAVNISPPSESDPVTILAENVPP 20220  
QY 172 ----- 171  
Db 20221 RIDLSVAMKSLITVKAGTNVCLDVTFGKPMPTVSWKKDGTLLKPAEGIKMAMORNLC TL 20280  
QY 172 ----- 171  
Db 20281 ELFSVNRKDSGYTITAENSSGKSATIKLVLDKPGPASVKINKMYSDRAMLSWEPL 20340  
QY 172 ----- 171  
Db 20341 EDGSGEITNIVDKRETSRPNWAQVSATVPITSCVSEKLIIEGHEYQFICAENKYGVDGP 20400  
QY 172 ----- 171  
Db 20401 VFTEPAIAKNPDPPGRCDPPVISNITKDHTVSWKPPADDGSGSPITGYLLEKRETOAVN 20460  
QY 172 ----- 171  
Db 20461 WTKVNRKPIIERTLKATGLQEGTEYEFRTAINKAGKPGFDASKAAYARDPQYPPAPPA 20520  
QY 172 ----- 171  
Db 20521 FPKVYDTRSSVLSWGKPAYDGGSPIIIGYLVEYKRADSDNWVRNCLPNLQKTRFEVTG 20580  
QY 172 ----- 178  
Db 20581 LMEDTOYQFRVYAVNKIGYSDPSDVPDKHYPKDILIPPEGELDADLRKTLILRAGVTMRL 20640  
QY 179 ----- 178  
Db 20641 YVPVGRPPKITSKPNVNLDRIGLDIKSTDFTDLRCENVNKYDAGKYLTLNLSG 20700  
QY 179 ----- 178  
Db 20701 KKEYTIVVKVLDTPGPPVNTVKEISKDSAYVTWEPPIIDGSPIIINVVQKRAERKSW 20760  
QY 179 ----- 178  
Db 20761 STVTTECSKTSFRVANLEBCKSYFFRVAENEYIGDGETRDAVKASQTPGPPVVDLKV 20820  
QY 179 ----- 178

Db 20821 SVSKSSCSIGMKPHSDGSGRIIGYVVDFLTENKQWVMKSLSIQYSAKDLTEGKEYTF 20880  
QY 179 ----- 178  
Db 20881 RVSAENENGECTPSEITVVARDDVVAPDLDLKGLPDLCYLAKENSFRKLTPIKKGKAPS 20940  
QY 179 ----- 178  
Db 20941 VSWKKGEDPLATDTRVSVESSAVNTTLLIVYDCQKSDACKYTTILKNVAGTKEGTISIKVV 21000  
QY 179 ----- 178  
Db 21001 GKPGIPTPIKFDEVTAEAMTLKWAPPKDDGSGEITNILEKRDVNNKWVTCASAVQKT 21060  
QY 179 ----- 178  
Db 21061 TFRVTRLHEGMEYTFRVSANENKYGVGELKSEPIVARHPFDPDAPPPNIVDVRHDSVS 21120  
QY 179 ----- 178  
Db 21121 LTWTDPKTKGSPITGYHLEPKERNSLWKANKTPIRMDFKVTGLTEGLEYSFRVMAI 21180  
QY 179 ----- 178  
Db 21181 NLAGVGKPSLPSEPVWALDPIPPCKPEVINITRNSVTLIWTEPKYDGGHKLGTGVIVEKR 21240  
QY 179 ----- 178  
Db 21241 DLPSKSMKANHVNPCEAFTVTDLVEGGKYEFRAIKANTAGIASPSESTETIICKDEY 21300  
QY 179 ----- 178  
Db 21301 EAPTIVLDPIKDLTIKAGDTIVLNAISILCKPLPKSSWKAGDIRPSDITQITSTPT 21360  
QY 179 ----- 178  
Db 21361 SSMLTIKYATRKADAGEYTTATNPGTKVHVKVTVLDVPGPPVEISNVSAEKATLTW 21420  
QY 179 ----- 178  
Db 21421 TPLEDGGSPIKSYILEKRETSRLLTWTVVSEDIQSCRHVATKLIQNEIYIFRSVAVNHG 21480  
QY 179 ----- 178  
Db 21481 KGPVQSEPVKMDRFGPPGPEKPEVSNVTKNATATVSWKRPVDDGSGEITGYHVERREK 21540  
QY 179 ----- 178  
Db 21541 KSLRWVRAIKTPVSDLRCKVTGLQEGSTYEFPSAENRAGIPPEASDSVLMKDAAAYPP 21600  
QY 179 ----- 178  
Db 21601 GPPSNPHVTDTTKKSASLAWGKPHYDGGLEITGYVVEHQYDEAWIKDITGTALRITQF 21660  
QY 179 ----- 178  
Db 21661 VVPDLQTKENYFRISAINDAGCEPAVDPVEIVEREMAPDFELDAELRRTLVRAGLS 21720  
QY 179 ----- 178  
Db 21721 IRIEVPKGRPAEVTWTKDNIKNRANIENTESFTLLIIEPCNRVDTGKFMVTIENPA 21780  
QY 179 ----- 178  
Db 21781 GKSGFVNVRVLDTPGPVNLNRPTDITKDSVTLHWDLPIDGSRITNIVIVEKRETRKS 21840  
QY 179 ----- 178  
Db 21841 YSTATTKCHKCTYKVTGLSEGCYFFRVMADENYIGEPTEPTEPVRKASEAPSPDLSNI 21900  
QY 179 ----- 178

Db 21901 MDITKSTVSLAWPKPKHDGSKITGYVIEAQRKSDQWTHITVTKGLECVVRNLTEGEY 21960  
QY 179 ----- 178  
Db 21961 TFOVMVNSAGRSAPRESRPVIVKEOTMLPELDLRGIYOKLVIAKAGDNKIVEIPVLRP 22020  
QY 179 ----- 178  
Db 22021 KPTVTWKGDQILKQORVNFETTATSTILNINECVSRSDGPPYPLTARNIVGEVGDVII 22080  
QY 179 ----- 178  
Db 22081 QVHDIPGPTGPIKFDEVSDFVFSWDPPENDGGVPISNYVEMRQDSTTIVELATV 22140  
QY 179 ----- 178  
Db 22141 IRTTYKATRLTGLEYOFRVKAQRNYGVGPGITSACIVANYFEKVPGPPTQVTAATKD 22200  
QY 179 ----- 178  
Db 22201 SMTISWHEPLSDGGSPILGYHVERKERNGLMQTVSKALVPGNIFKSSGLTDGIAEYFRV 22260  
QY 179 ----- 178  
Db 22261 IAENMAGSKPSKPSEPMALDPIDPPGKVPPLNITRHTVTTLKWAPEYTGFKITSYIV 22320  
QY 179 ----- 178  
Db 22321 EKRDLPNGRLKANFSNILENETVSGLTEDAAYEFRTAKNAAGAI SPPSEPSDAITCR 22380  
QY 179 ----- 185  
Db 22381 DOVEAPKIKVDVKFKDTVLKAGAFRLADYSGRPPTMWSKDGKEGTAKLEIKTA 22440  
QY 186 ----- 185  
Db 22441 DFSTNLVNDSTRSGAYTLTATNPGGAKHIFNVKVLDRPGPEGLAVTEVTSEKV 22500  
QY 186 ----- 185  
Db 22501 LSWFPLDDGAKIDHYIVOKRETSRLAWTNVASEVQVTKLVTKLLKGNEYIFRVMVN 22560  
QY 186 ----- 185  
Db 22561 KYGVGELESEPVAVNPYGPDPKPNPEVTTITKDSMVVVCWGHPSDGGSEIINIVYER 22620  
QY 186 ----- 185  
Db 22621 RDKAGORWIKCNKKTLDLRYKVSGLTEGHEYEFRI MAENAGISAPSPSPFYKACDTV 22680  
QY 186 ----- 185  
Db 22681 FRPGPGNPRVLDTSRSSISIAWNKPIYDGSSEITGYWVEIALPEDEWQIVTPPAGLKA 22740  
QY 186 ----- 185  
Db 22741 TSYTITGLTENOYKIRIYAMNSEGLGEPALVPGTPKAEDRMLPPEIELDADRKVVTIR 22800  
QY 186 ----- 185  
Db 22801 ACCTLRLEFVPIKGRPAPEVKWARDHGESLDKASIESTSYTLLIVGNVNRFDGKYILTV 22860  
QY 186 ----- 185  
Db 22861 ENSSGSKSAFNVVRVLDTPGPPQDLKVKEVTKSVTLTWDPPLDGGSKIKNYIVEKRES 22920  
QY 186 ----- 185  
Db 22921 TRKAYSTVATNCHKTWKVDQLQEGCSYVFRVLAENEYIGLPAETAESVKASERPLPG 22980  
QY 186 ----- 185  
Db 22981 KTTLMDVTRNSVLSWEKPEHDDGSRILGYIVEMQTKSGDKWATCATVKVTEATITGLIQ 23040

QY 186 ----- 185  
Db 23041 GEYSRVSQAQNEKGISDPRQLSPVIAKDLVIPPAFKLLFNTFTVLAGEDLKVDVFFTG 23100  
QY 186 ----- 185  
Db 23101 RPTPAVTHKDNVPLKQOTTRVNAESTENNLLTIKDACREDVGHYVVKLTNSAGEAIEL 23160  
QY 186 ----- 185  
Db 23161 NVIVLDKPGPPTGPVKMDEVTA DSITLSWGPYPKYDGGSSINNYIVEKRDSTTTWQIVSA 23220  
QY 186 ----- 185  
Db 23221 TVARTTIKACRLKTGCEYQFRIA AENRYCKSTYLNSEPTVAQYPPKVPGPPTPVVTLSS 23280  
QY 186 ----- 185  
Db 23281 RDSMEVQWNEPISDGGSRVIGYHLERKERNLSILWKLNKTPIPQTKFKTTGLEEGVEYEF 23340  
QY 186 ----- 185  
Db 23341 RVSAENIVGIGKPSKVSECYVARDPCDPPRPEAIIVTRNSVTLOWKKPTYDGGSKITCY 23400  
QY 186 ----- 185  
Db 23401 IVEKKELPEGRWMAKASFTNIIDTHFEVTGLVEDHRYEFVRVIA RNAAGVSESESTGAI 23460  
QY 186 ----- 185  
Db 23461 ARDEVDPPIRISMDPKYKDTIVVHAGESFKVDADIYKPIPTIOWIKGDOELSNARLEIK 23520  
QY 186 ----- 185  
Db 23521 STDFATSLSVKDAVRVDGSGNYILKAKNAVAGERSVTNVKVLDRPPEGPVVISGVTAEK 23580  
QY 186 ----- 185  
Db 23581 CTLANKPPLQDGGSDIINIVYERRETSRLVWTVVDANVQTLSCVTKLLEGNEYTRINA 23640  
QY 186 ----- 185  
Db 23641 VNKYGVGELESEPVVAKNPFVVPDAPKAPETVTTKDSMIVVWBERPASDGGSEILGYVL 23700  
QY 186 ----- 185  
Db 23701 EKRDKEGIRWTRCHRRLLIGELRLRVTGLIENHDYEFVSAENAGLSEPPSPSAIYOKACD 23760  
QY 186 ----- 185  
Db 23761 PIYKPGPNPKVIDITRSSFVLSWSKPIYDGCCEIQGYIVEKCDVSGEWMTCTPTTCI 23820  
QY 186 ----- 185  
Db 23821 NKTNIEVEKLEKHEYNFICAINKAGVHADVPGPPIIVEKLEAPDIDLDELRLKIIN 23880  
QY 186 ----- 185  
Db 23881 IRAGSLRLFVPIKGRPTEPVKWGVKVDGEIRDAAIIDVTSSFTSLVLDNVNRYDSGKYTL 23940  
QY 186 ----- 185  
Db 23941 TLENSSGTKSAFVTVRVLDTPSPPVNLKVTEITKDSVSTIWEPPLLDGGSKIKNYIVEKR 24000  
QY 186 ----- 185  
Db 24001 EATRKSAAVVTNCHKNSWKIDQLQEGCSYVFRVTAENEYIGLPAQTADPIKVAEVP 24060  
QY 186 ----- 185  
Db 24061 PGKITVDVTRNSVLSWTKPEHDDGSKIYIVEMQAHSKSEKARVKSQAIVITNL 24120

QY 186 ----- 185  
Db 24121 TQGEVLFVRVAYNEGRSDPRSLAVPIVAKDLVIEPDVKPAPSSYVQVGQDLKIEVPI 24180  
QY 186 ----- 185  
Db 24181 SGRPKPTITWKDGLPLKOTTRINVTDSLDTLSIKETHKDDGGYGITVANVVGKTA 24240  
QY 186 ----- 185  
Db 24241 SIEIVLDRPKPPKGVKFDVSAESITLSWNPPLYTGGCQITNYIVQKRDTTITVWDVV 24300  
QY 186 ----- 185  
Db 24301 SATVARTTLKVTKLKTGTEYQFRIFAENRYGQSFALESPIVAQYPYKBPPTPFATA 24360  
QY 186 ----- 185  
Db 24361 ISKDSMVIOMHEPVNNGSPVIGYHLERKERNISILWTKVNTLIHDTQFKAQNLBEGIEY 24420  
QY 186 ----- 185  
Db 24421 EFRVYAENIVGKASKNSECYVARDPCDPPGTPPEIMVKRNEITLQWTKPVYDGGSMIT 24480  
QY 186 ----- 185  
Db 24481 GYIVEKRDLPDGRMKASFTNVLETQFTVSGLTEDQRYEYFVIAKNAAGAIKSPSDTGP 24540  
QY 186 ----- 185  
Db 24541 ITAKDEVLPRIKSMDFKRDITVWNAGETFRLEADVHGKPLPTIEWLRGDKETEESARCE 24600  
QY 186 ----- 185  
Db 24601 IKNTDFKALLIVKADRIDGGVILRASNVAGSKSPVNVKVLDRPPEGPVQVTGVS 24660  
QY 186 ----- 185  
Db 24661 EKCSLTWSPLODGGSDISHYVVEKRETSRLAWTVVASEVVTNSLKVTKLLEGNEYVFRI 24720  
QY 186 ----- 185  
Db 24721 MAVNKYGVGEPLSAPVLMKNPFLPGPKPLEVTNIAKDSMTVCWNRPSDGGSEIIGY 24780  
QY 186 ----- 185  
Db 24781 IVEKDRSGIRWIKCNKRRTIDRLRVTGLTEDHEYEFVRSNAAGVGEPSPATVYYKA 24840  
QY 186 ----- 185  
Db 24841 CDPVFKPGPPTNAHIVDTTKNSITLAWGKPIYDGGSEILGYVEICKADEEWOI VTPQT 24900  
QY 186 ----- 185  
Db 24901 GLRVTRFEISKTEHOEYKIRVCALNKVGLGEATSVPGTVKPEDKLEAPELDDLSELKRG 24960  
QY 186 ----- 185  
Db 24961 IIVRAGGSARIHIPFKGRPTPEITWSREGEFTDKVQIEKGVNYTQLSIDNCDNRDAGKY 25020  
QY 186 ----- 188  
Db 25021 ILKLENSGSKSAFVTVKVLDTGPPQNLAVKEVRKDSAFVWEPPIIDGAKVKVVID 25080  
QY 189 ----- 188  
Db 25081 KRESTRKAYANSSKSKTSFKVENLTEGAIYYFRVMAENEFVGVGVVETVDVAKAAEPP 25140  
QY 189 ----- 188  
Db 25141 SPPGKVTLTDVTSASLWKEPHEHGGSRVLGYVVEMQPKGTEKWSIVAESKVCNAVVT 25200  
QY 189 ----- 188

Db 25201 GLSSGOEQYQFRVKAYNEKGKSDPRVLGVPVIAKDLTIQPSLKLFPNTYSIQAGEDLKIEI 25260  
QY 189 ----- 188  
Db 25261 PVIGRPRPNISWVKDGEPLKQTTTRVNVEETATSTVLHIKEGNKDDFCYKTVTATNSAGTA 25320  
QY 189 ----- 188  
Db 25321 TENLSVIVLEKPPGPPVPRFDEVDSADFVVVISWEPPAYTGGCOISNVIIVEKRDTTTTTWH 25380  
QY 189 ----- 188  
Db 25381 MVSATVARTTIKTTLKTGTGYQFRIFAENRYGKSAPLDSKAVIVQYFPKPGPGTPEV 25440  
QY 189 ----- 188  
Db 25441 TSISKDQMLVQWHEPVNDGGTKIIGYHLEQEKNSILWKLNKTPIQDTKFKTTGLDEGL 25500  
QY 189 ----- 188  
Db 25501 EYEFKVSANIVGIGKPSKVSECFVARDPCDPPCRPEAIVITRNNVTLKWKPAYDGGSK 25560  
QY 189 ----- 188  
Db 25561 ITGYIIVEKKDLPDGRMKASFTNVLETEFTVSGLVEDQRYEYFVIARNAAGNFSEPSDSS 25620  
QY 189 ----- 188  
Db 25621 GAITARDEIDAPNASLDPKYKDVIVVHAGETFVLEADIRGKPIPDVVWMSKDGKELEETA 25680  
QY 189 ----- 188  
Db 25681 RMEIKSTIOKTTLVVVKDCIRTDGGQYILKLSNVGGTKSPITVKVLDLRPGPPGGLKVTG 25740  
QY 189 ----- 193  
Db 25741 VTAECYLAWNPPLODGGANISHYIIEKRETSRLSWTQSTVEQALNKYKTVKLLPGNEYI 25800  
QY 194 ----- 193  
Db 25801 FRVMAYNKYGIGEPLESGPVACNPKYKPPGPPSTPEVSATTKDSMVVTVWARPVDDGGTEI 25860  
QY 194 ----- 193  
Db 25861 EGYILEKRDEKGYRWTCKNKKTLTDLRLRVTGLTEGHSYEFVRAAENAGVGEPSPSVF 25920  
QY 194 ----- 193  
Db 25921 YRACDALYPPGPPSPNKVTDTSRSSVSLAWSKPIYDGGAPVKGYVVEVKEAADEWTTCT 25980  
QY 194 ----- 193  
Db 25981 PPTGLOGKQFTVTKLKENTYENFRICAINSEGVGEPATLPGSVVAQERIEPPEIELDADL 26040  
QY 194 ----- 193  
Db 26041 RKVVVLRASATLRLFTVIKGRPEPEVKWEKAEGLITDRAQIEVTSSTFMLVIDNVTRFDS 26100  
QY 194 ----- 193  
Db 26101 GRYNLTLENNSGSKTAFVNVVRLDPSAPVNLTIREVKKDSVTLISWEPPLIDGGAKITNY 26160  
QY 194 ----- 193  
Db 26161 IVEKRETRKAYATITNNCTKTTFRLENLOEGCSYFVRVLASNEYGIGLPAETTEPVKVS 26220  
QY 194 ----- 193  
Db 26221 EPPLPGRVTLVDVTRNTATIKWEKPESDGGSKITGVVEMQTKGSEKWSCTQVKTLEA 26280  
QY 194 ----- 193

Db 26281 TISGLTAGEEYVFRVAANNEKGRSDPROLGVPIIARDIEIKPSVSLPFFHTFNVKAREQLK 26340  
Qy 194 ----- 193  
Db 26341 IDVPFKGRQATVNRKDGQTLKETTRVNSSKIVTSLIKEASKEDVGIELCVNSNA 26400  
Qy 194 ----- 193  
Db 26401 GSITVPIIIVLDRPGPGPIRIDEVSCDSITISWNPPEYDGGCOISNVIVEKETTSTT 26460  
Qy 194 ----- 193  
Db 26461 WHIVSOAVARTSIKIVRLTTGSEYQFRVAENRYCKSSYESSAYVAEYFPFPGPGPT 26520  
Qy 194 ----- 193  
Db 26521 KVVHATKMTLVWQVPVNDGSGSRVIGHLEYKERSILWSKANKILLIADTOMKVSGLDE 26580  
Qy 194 ----- 193  
Db 26581 GLMYEYRVVAENIAGIGKCKSCPEVPARDPCDPGQPEVTNITRKSYSKLWCKPHYDGG 26640  
Qy 194 ----- 193  
Db 26641 AKITGYIVERRELDPGRWLKCNITNIQETYFEVTELTEDQRYEYRFAARNAADSVSEP 26700  
Qy 194 ----- 193  
Db 26701 STGPIIVKDDVEPRVMDVFRDVIIVKAGEVLKINADIAGRPLPVISWAKDGEIER 26760  
Qy 194 ----- 193  
Db 26761 ARTEIISTDNHTLLTVKDCIIRDTQGYVLTKNVAGTRSAVAVCKVLDKPGPPAGPLEIN 26820  
Qy 194 ----- 193  
Db 26821 GLTAEKCSLSWGRPOEDGADIDYIIVEKRETSHLAWTICEGELQMTCKVTLLKGNEY 26880  
Qy 194 ----- 193  
Db 26881 IFRVTGVNKYVGGEPLSVAIKALDPFTVPSPPTSLEITSVTKESMTLCWSRPSDGGSE 26940  
Qy 194 ----- 193  
Db 26941 ISGYIIRERKSLRWVRNKKPVYDLRVKSTGLREGCEYEVRYVAENAGLSLPSETSP 27000  
Qy 194 ----- 193  
Db 27001 LIRADPVLPSPPSKPIVDGKTTITIAWVKPLFDGGAPITGYTVEYKKSDDTDWKTS 27060  
Qy 194 ----- 193  
Db 27061 IQSLRGTEYITISGLTTGAEYVFRVSKVNVKASDPSDSDPOIAKEREELPFDIDSEMR 27120  
Qy 194 ----- 193  
Db 27121 KTLIVKAGASFTMTVPFRGRPVPNVLSKPDOLTRAYVDTDSRTSLTIENANRNDG 27180  
Qy 194 ----- 193  
Db 27181 KYTLIQNLVSAASLTLLVVKVLDTPGPPTNITVQDVTKEASAVLSWDVPENDGGAPVKNYH 27240  
Qy 194 ----- 193  
Db 27241 IEKREASKAMVSVTNNCNRLSYKTNLQEGAIYFRVSGENEFVGVGIPAEIKEGVITE 27300  
Qy 194 ----- 193  
Db 27301 KPSPEKLGVTISKSDSVSLTWLKEPHDGGSRIVHYVVEALEKGOKNWKCAVAKSTHHV 27360  
Qy 194 ----- 193  
Db 27361 VSGLRSENSEYFRFAENQAGLSDPRELLLPVLVLIKEQLEPPEDMKNFPSTHYVVRAGSN 27420

Qy 194 ----- 193  
Db 27421 LKVIDIPISGKPLPKVTLSRDGVPLKATMRFNTEITAENLTINLKESVTADAGRYEITAA 27480  
Qy 194 ----- 193  
Db 27481 SSGTTKAFINIVLDRPGPGTPPVVISDITEBSVTLKWEPPKYDGGSOVTNVLKRETS 27540  
Qy 194 ----- 193  
Db 27541 TAVMTEVSATVARTMMKVKMLTTGEEYQFRIKAENRFGISDHIDSACVTVKLPYTPGPP 27600  
Qy 194 ----- 193  
Db 27601 STPWTVNTRESITVGVWHEPVNGSGSAVGVHLEMKDRNSILWQKANKLVIRTHFKVTT 27660  
Qy 194 ----- 193  
Db 27661 ISAGLIYEFYVAENACGVGKPSHPSEVLAIDACEPPNRVIRITDISKNSVLSWQOPAF 27720  
Qy 194 ----- 193  
Db 27721 DGSKITGYIVERRDLDPGRWTKASFNTVETQFIISGLTQNSQYEFYRFAARNAVSGISN 27780  
Qy 194 ----- 193  
Db 27781 PSEVVGPICTIDSYGGPVLDLPLEYTEVVKYRAGTSVKLRAGISGKPAPTIEWYKDDKEL 27840  
Qy 194 ----- 193  
Db 27841 QTNALVCVENTDLASILIKDADRNLSCYELKLRNANGSASATIRVOILDKPPGGPI 27900  
Qy 194 ----- 193  
Db 27901 EFKVTAEKITLLWRPPADGGAKITHYIVEKRETSRVVSMVSEHLEECITTTKIIG 27960  
Qy 194 ----- 193  
Db 27961 NEYIFRVAVNKYIGIGEPLSDSVVAKNAFVTPGPPGPIEVTKITKNSMTVWSRPIADG 28020  
Qy 194 ----- 193  
Db 28021 GSDISGYFLEKRDKSLGWFKVLKETIRDTROKVTGLTENSQYRVCAVNAAGQGPSE 28080  
Qy 194 ----- 193  
Db 28081 PSEFYKAADPIDPPGPPAKIRIADSTKSSITLWWSKPVYDGGSAVTVYVVEIROGEEEW 28140  
Qy 194 ----- 193  
Db 28141 TTVSTKGEVRTTEYVVSNLKPGVNYFRVSAVNCAGOGEPTEMNEPVQAKDILEAPEIDL 28200  
Qy 194 ----- 193  
Db 28201 DVALRTSVIAKAGEDVOVLIPFKGRPPPTVTRKDEKNLGSDARYSTIENTDSSLLTTPQ 28260  
Qy 194 ----- 193  
Db 28261 VTRNDTGKYLTIENGVEPKSSTSVKVLDTPAACQKLOVKHVSRGTVTLMLDPPPLIDG 28320  
Qy 194 ----- 193  
Db 28321 GSPINIYVIEKRDATKRTWSVVSHKCSSTSPKLDLSEKTPFFFRVLAENEIGIGECET 28380  
Qy 194 ----- 193  
Db 28381 TEPVKAEEVPAPIRDLNMDSTKTSVILSWTKPDFDGGSVITEYVVERKKGEGTWSHAG 28440  
Qy 194 ----- 193  
Db 28441 ISKTEIEVSQLEQSVLEFRVFAKNEKGLSDPVTIGITVTKELIITPEVDLSIDPGAQV 28500

Qy 194 ----- 193  
Db 28501 TVRIGHNVHLEPYKGRPKPSISMLKDLPLKESEFYRFSKTENKIILSIKNAKEHGGK 28560  
Qy 194 ----- 193  
Db 28561 YTVILDNAVCRIAPITVITLGPSPKPKPIRFDEIKADSILSWDPEDNGGGEITCYS 28620  
Qy 194 ----- 193  
Db 28621 IEKRETSOTNMKMCSSVARTTFKVPNLVKADEYQFRVRAENRYGVSOPLVSSIIVAKHQ 28680  
Qy 194 ----- 193  
Db 28681 FRIPGPGKPYIYNVTSQMSLTWDAPVYDGGSEVTGFHVEKKERNSTLWKVNTSPISG 28740  
Qy 194 ----- 193  
Db 28741 REYRATGLVGLDYQFRVYAENSAGLSSPSPSKFTLAVSPVDPGTPDYIDVTRITL 28800  
Qy 194 ----- 193  
Db 28801 KWNPLRDGSGKIYGVYSTIEKKOGRNWRVRCNFTDVSECQYTVTGLSPGDRYEFRIIARNA 28860  
Qy 194 ----- ppv----- 196  
Db 28861 VGTISPPSSGSIIMTRDENVPPIVEFGPEYFDGLIIKSGESLRIKALVQGRPVPRVTFW 28920  
Qy 197 ----- 196  
Db 28921 KGVIEIKRMNMEITDVLGSLFVRDATRDHRGVYTVVEAKNAGSAAEKVKVQDTPG 28980  
Qy 197 ----- 196  
Db 28981 KVGPIRFTNITGEKMTLWMDAPLNDGCAPITHYIIIEKRETSRLAWALIEDKCEAQSYTA 29040  
Qy 197 ----- LDS----- 200  
Db 29041 IKLINGNEYQFRVSAVNFVGRPLDSDPVVAQIQYTPDAPGPEPSNITGNSITLWA 29100  
Qy 201 ----- 200  
Db 29101 RPESDGGSEIOQYILERREKSTRWVKVISKRPISETRFRKVTGLTEGNEYEFHVMAENAA 29160  
Qy 201 ----- 200  
Db 29161 GVGPSGISRLIKCREPVNPPGPTTVVKVTDTSKTTVSLENSKPVFDGGMELIIGYIEMC 29220  
Qy 201 ----- 200  
Db 29221 KADLGDWHKVNAAECVKTRYVTDLQAGEEYKFRVSAINGAGKGDSCVETGTIKAVDRLT 29280  
Qy 201 ----- 200  
Db 29281 APELDIDANFKQTHVVRAGASIRLFYAYQGRPTTAVWSKPDNSLNRADITHHDDSFTL 29340  
Qy 201 ----- 200  
Db 29341 TVENCNRNDAGKYTLTVENNSGSKSIITFVKVLDTPGPPGPIITFKDVTGRSATLWMDAPL 29400  
Qy 201 ----- 200  
Db 29401 LDGGARIHHYVVEKREASRSQWVISEKTRQIFKVNDAEAGVPPYFRVSAVNEYGVGEP 29460  
Qy 201 ----- 200  
Db 29461 YEMPEPIVATEQAPAPRRRLDVVDTSKSSAVLAWLKPDDHGGSRITGYLLEMROKGSDFW 29520  
Qy 201 ----- 200  
Db 29521 EAGHTKQLTFTVERLVEKTEYFRVKAENDAGYSEPREAFSSVIIKEPOIEPTADLTGIT 29580  
Qy 201 ----- 200

Db 29581 NQLITCKAGSPFTIDVPISGRPAKPVTKLEMLKETDRYSITTTKORTTLTVKDSMRG 29640  
Qy 201 --GSFFL----- 205  
Db 29641 DSGRYFLTLENTAGVKTFSVTVVVIGRPGVPTGPIEVSSVSAESCVLWSGPKDGGGTEI 29700  
Qy 206 ----- 205  
Db 29701 TNYIVEKRESCTTAWQLVNSSVKRTOIKVTHLTMYEYSFRVSSENRFGVSKPLESAPII 29760  
Qy 206 ----- 205  
Db 29761 AEHPFVPPSAPTREPVYHVSANAMSIRWEEPYHDGGSKILGYWVEKKERNITLWKENKV 29820  
Qy 206 ----- 205  
Db 29821 PCLECNKYKVTGLVEGLEVFQFRTYALNAAGVSKASEASRPIAQNPDVDPAGRPVTDVTRS 29880  
Qy 206 ----- 205  
Db 29881 TVSLINAPAYDGGSKVVGYYIIERKPVSEVGDGRWLCNYYTIVSDNFFTVTALSEGDTYE 29940  
Qy 206 ----- 205  
Db 29941 FRVLAKNAAGVISKGSESTGPTCRDEYAPPKAELDARLHGLDVTIRAGSDLVLDAAVGG 30000  
Qy 206 ----- 205  
Db 30001 KPEPKIIWTGDKELDCEKVSLOYTGKRATAVIKFCDRSDSGKYTLTVKNASGTFKAVSV 30060  
Qy 206 ----- 205  
Db 30061 MKVLDSPGCGKLTVSRTQEKTLAWSLPQEDGGAETHYIVERRETSRLNWNVIVEGE 30120  
Qy 206 ----- 205  
Db 30121 CPTLSYVVTRLIKNNEYIFRVRANVKNYGPVVESEPIVARNSTFIPSPGPIEVEGTGK 30180  
Qy 206 ----- 205  
Db 30181 EHIIQWTKPESDGGNEISNLYLDKREKSLRWTRVNKYVYDTRLKVTLMEGCDYOF 30240  
Qy 206 ----- 205  
Db 30241 RVTAVNAGNSEPSEASNFISCREPSYTPGPPSAPRVYDTHKHSISLAWTKPMYDGGTDI 30300  
Qy 206 ----- 205  
Db 30301 VGYVLEMQEKTDQWYRVHTNATIRNTEFTVPLKMGOKYSFRVAANVKNMGSEYSESIA 30360  
Qy 206 ----- 205  
Db 30361 EIEPVERIEIPDELAADDLKKTVTIRAGASLRLMVSVSRPPPVITWSKQIDLASRAII 30420  
Qy 206 ----YSKLTVDK----- 213  
Db 30421 DTTESYLLIYDKNRYDAGKYTIEAENQSGKKSATVLVKYDTPGPCPSVKVKEVS RDS 30480  
Qy 214 ----- 213  
Db 30481 VTITWEIPTIDGAPVNNYIVEKREAAAMRAFKTVTKCKTLYRISGLVEGTMYYFRVLP 30540  
Qy 214 ----- 213  
Db 30541 ENIYGICEPCTSDAVLSEVPLVPAKLEVVDVTKSTVTLAWEKPLYDGGSRLTGYVLEA 30600  
Qy 214 ----- 213  
Db 30601 CKAGTERMKVVTLKPTVLEHTVTSLEGEQYLFRIRAONEKGVSEPRETVTAVTVQDLR 30660  
Qy 214 ----- 213

Db 30661 VLPTIDLSTMPQKTIHVPAGRPVELVPIAGRPPPAASWFFAGSKLRESERVTVETHKV 30720  
QY 214 ----- 213  
Db 30721 AKLTIRETTIRDTGEYTLKKNVTGTTSETIKVILDKPGPPGPIKIDEIDATSTISW 30780  
QY 214 ----- 213  
Db 30781 EPELDGAPLSGVVEQDAHRPGWLPVSESVTRSTFKFTRLTEGNEYVFRVAATNRF 30840  
QY 214 ----- 213  
Db 30841 IGSYLQSEVIECRSIRIPGPETLQIFDVSRDGMTLTWYPEDDGGSQVTGYIVERKEV 30900  
QY 214 ----- 213  
Db 30901 RADRWVRVNVKVPMTYRSTGLTEGLEYEHRVTAI NARGSGKPSRPSKPIVAMDPIAPP 30960  
QY 214 ----- 213  
Db 30961 GKQNPRTDTRTSVSLAWSVPEDEGGSKVTGYLIEMQVDOHEWTKCNTPTPKIREYT 31020  
QY 214 -----SRWQOG----- 219  
|:|:  
Db 31021 LTHLPOGAEYRFRVLACNAGGPGPAEVPGVTKVTEMLEYDPYELDERYQEGIFVRGGV 31080  
QY 220 ----- 219  
Db 31081 IRLTIPIKGPFFICKWKKEGQDISKRAMIATSETHTELVIKEADRGDSGTVDLVLENKC 31140  
QY 220 ----- 219  
Db 31141 GKAVYIKVRVIGSPNSPEGPLEYDDIQVRSVRVSRNPADGGADILGILYLERVPCA 31200  
QY 220 ----- 219  
Db 31201 AWYIDSRVRGTSLVVKGKLKENVEYHFRVSAENQFGISKPLKSEEPVTPKTLNPPPEPS 31260  
QY 220 ----- 219  
Db 31261 NPPEVLDTKSSVLSWSRPKDDGSRVVGYYIERKETSTDKWVRHNTQITTTMTYVTG 31320  
QY 220 ----- 219  
Db 31321 LVPDAEYQFRIIAQNDVGLSETSPASEVPVCKDPFDKPSQGELEILSISKDSVTLQWEK 31380  
QY 220 ----- 219  
Db 31381 PECDDGKEILGYWVEYRQSGDSAMKSNKERIKDKOFTTIGLLEATEYEFVRFAENETGL 31440  
QY 220 ----- 219  
Db 31441 SRPRRTAMSIKTLTSGEAPGIRKEMKDVTTKLGEAAQLSCQIVGRPLPDIKWYRFGKEL 31500  
QY 220 ----- 219  
Db 31501 IOSRKYKMSSDGRHTLTVMTEBEQDEGYVTCIATNEVGEVETSSKLLQATPQHPGPY 31560  
QY 220 ----- 219  
Db 31561 LKEKYGAVGSTLRHLHWYIGRPVPAMTWFGHKKLONSENITIENTEHTHLVMKNVOR 31620  
QY 220 -----NVF----- 222  
|:|:  
Db 31621 KTHAGRYKVOLSNVFGTVDAILDVEIQDKPDKPTGPTVIEALLKNSAVISWKPPADDGGS 31680  
QY 223 ----- 222  
Db 31681 WITNYVVERCEAKEGAEMOLVSSAISVTTCRIVNLNTENAGYFRVSAQNTFGISDPLEVS 31740  
QY 223 ----- 222  
Db 31741 SVVIKSPPEKPGAPGKPTITAVTKDSCVVAWKPASDGGAKIRNYYLEKREKKONKWIS 31800

QY 223 ----- 222  
Db 31801 VTTEEIRETVFSVKNLIEGLEVEFRVKCENLGSESEISEIPITPKSDVPIQAPHKEE 31860  
QY 223 ----- 222  
Db 31861 LRNLNVYQSNATLVCKVTGHPKPIVKWYRQGEKIIADGLKYRIQFKGYHOLIISVT 31920  
QY 223 ----- 222  
Db 31921 DDDATVYQVRATNOGGSVSGTASLEVEVPAKIHLPKLTLEGMAVHALRGEVWSIKIPFG 31980  
QY 223 ----- 222  
Db 31981 KPDPVITWQKGQDLIDNNHGYQVIVTRSTSLVFPNGVERKDGAGFYVVCANKNRFIDQKT 32040  
QY 223 ----- 222  
Db 32041 VELDVADVPDPRGVKVSVDVSRDSVNLTWTEPASDGSKITVYIVEKCATTAERWLRVGQ 32100  
QY 223 ----- 222  
Db 32101 ARETRYVINLFCKTSYQFRVIAENKFGLSKPSSEPTITTKEDKTRAMNYDEEVDRE 32160  
QY 223 ----- 222  
Db 32161 VSMTKASHSTKELYEKYMAIEDLGRGEGFIVHRCVETSSKKTYMAKFKVKGTDQVLVK 32220  
QY 223 ----- 222  
Db 32221 KEISILNIAHRNHLHESFESMEELVMFEFISGLDIFERINTSAFELNEREIVSYVH 32280  
QY 223 ----- 222  
Db 32281 QVCEALQFLHSHNIGHDIRPENIIYQTRRSSTIKIIEFGARQLKPGDNFRLLFTAPEY 32340  
QY 223 ----- 222  
Db 32341 YAPEVHQHDVSTATDMNSLGLTVVLLSGINPFLAETNQIIENIMNAEYTFDEEAPKE 32400  
QY 223 ----- 222  
Db 32401 ISIEAMDFVDRLLVKERKSRMTASEALQHPWLKQKIERVSTKVIRTLKHRRYYHTLIK 32460  
QY 223 ----- 222  
Db 32461 LNMVYSAARISCGGAIRSQKGSVAKVAVASIEIGPVSQIMHAVGEGGHVYCKIEN 32520  
QY 223 ----- 222  
Db 32521 YDOSTQVTVYFGVRQLENSEKYEITYEDGVAILYVKDITKLDGTYRCKVVDYGEDSSY 32580  
QY 223 ----- 222  
Db 32581 AELFKVGREVYDYICRRTMKKIKRRDTMTLLRPPETPLPLYNKATYGVNRFVGTI 32640  
QY 223 ----- 222  
Db 32641 TVHPEPHVTWYKSGOKIKPGDNDKKYTFESDKGLYQLFTINSVTTDDAEYTVVARNKYGE 32700  
QY 223 -SC-  
|:|:  
Db 32701 DSCAKLAVTLHPPTDSTLRPMFKRLLANABEQSQSCVCFIRYSGIPPPPLKWKEDQ 32760  
QY 225 -----SVMEAL----- 231  
|:|:  
Db 32761 PLSLGPNIETIIHEGLDYVALHIRDTLPEDTGYIRVATATAGTSCOAHLQVERLRYKKQ 32820  
QY 232 ----- 231  
Db 32821 EFKSKEEHERHVQKOIDKTLRMAEILSGTESVPLTQVAKEALREAAVLYKPAVSTKTVKG 32880



QY 232 ----- 231  
Db 32881 EFRLEIEKKEKRLMPYDVPEPRKYQFTIEDQRIKQFVPMQDMKWKYKIRDOYEMP 32940  
QY 232 ----- 231  
Db 32941 GKLDVVQKPKRIRLSRWQFYVPLPRITDQYRKPWRIPKLSQDDLEIVRPARRTPS 33000  
QY 232 ----- 231  
Db 33001 PDYDFYRPRRSLGDISDELLPIDDYLA MKRTEERLRLEEELELGSASPPSRSP 33060  
QY 232 ----- 231  
Db 33061 HFEISSLRYSSQAHVKVEETRKDFRYSTVHIPTKAEASTSYAELRERHAQAAYRQPKOR 33120  
QY 232 ----- 231  
Db 33121 ORIMAEREDELLRPVTTTQHLSEYKSELD FMSKEEKSRRQRREVTEITEIEEYEI 33180  
QY 232 ----- 231  
Db 33181 SKHAQRESSSSASLLRRRLSFTYIELMRPVSELIRSPQPAEYEDDTERSPPTPER 33240  
QY 232 ----- 231  
Db 33241 TRPRSPVSPSSRSRFSARFIDFSRYESMKAALKTKTOKSERKYEVLSSQPFLLDHA 33300  
QY 232 ----- 233  
Db 33301 PRITLRMRSHRVP CGONTREFILNVQSKPTAEVKYHNGVELQSSKIHNTSGVLTLEI 33360  
QY 234 ----- 233  
Db 33361 LDCHTDSDGYRAVCNYKGEASDYATLDVTGGDYTYASQRDEEVPRSVFFELRTEA 33420  
QY 234 ----- 233  
Db 33421 YAVSSFKKTSMEASSVREVKSQMTETRESLSYHSASAEKSALEEKSLSEKSTTR 33480  
QY 234 ----- 233  
Db 33481 KIKTTLAARILTKPRSMTVYEGSARFSCDTPGVPVTWLRKGQVLSARHQVTTK 33540  
QY 234 ----- 233  
Db 33541 YKSTFEISSVQASDEGNYSVVENSEGKQAEFTLTIQKARVTEKAVTSPPRVKSPEPRV 33600  
QY 234 ----- 233  
Db 33601 KSPEAVKSPKRVKSPEPSHPKAVSPPTETKPTPEKVQHLVPSAPPKITQFLKAEASKEIA 33660  
QY 234 ----- 235  
Db 33661 KLTCWESSVLRAKEVTWYKDGKKLKGNGHFQFHSADGTIELKNNLTSQDQGYVCEI 33720  
QY 236 ----- 235  
Db 33721 SGEIGTSKTNLQPMQAFKSIHEKSVKISETKKSDQKTTESTVTRKTEPKAPEISSKPV 33780  
QY 236 ----- 235  
Db 33781 IVTGLQDVTYSSDVAKFAVKATGEPRPTAIWKDGKAITQGGKYKLSEDKGFFLEIHK 33840  
QY 236 -----  
TQKSLSLSPK 246  
|||: :||  
Db 33841 TDTSDGLYCTVKNKSAGSVSSCKLTIKAIKDTQAKQVSTQKTSEITPQK 33891  
RESULT 14  
Q10466  
ID Q10466 PRELIMINARY; PRT: 26926 AA.  
AC Q10466;

DT 01-NOV-1996 (TREMBLrel. 01, Created)  
DT 01-NOV-1996 (TREMBLrel. 01, Last sequence update)  
DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)  
DE Titin, heart isoform N2-B (EC 2.7.1.-) (Connectin).  
OS Homo sapiens (human).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
OX NCBI\_TaxID=9606;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC TISSUE=HEART;  
RX MEDLINE=96026330; PubMed=7569978;  
RA Labeit S.; Kolmer B.;  
RT "Titins: giant proteins in charge of muscle ultrastructure and  
RT elasticity";  
RL Science 270:293-296(1995).  
RN [2]  
RP SEQUENCE OF 2277-25376 FROM N.A.  
RX MEDLINE=92258380; PubMed=1582406;  
RA Labeit S.; Gautel M.; Lakey A.; Trinick J.;  
RT "Towards a molecular understanding of titin";  
RL EMBO J. 11:1711-1716(1992).  
RN [3]  
RP SEQUENCE OF 1976-2014 FROM N.A.  
RA Labeit S.;  
RL Submitted (DEC-1994) to the EMBL/GenBank/DBJ databases.  
RN [4]  
RP CHARACTERIZATION.  
RX MEDLINE=95331314; PubMed=7607248;  
RA Gautel M.; Castiglione-Morelli M.A.; Pfuhl M.; Motta A.; Pastore A.;  
RT "A calmodulin-binding sequence in the C-terminus of human cardiac  
RT titin kinase";  
RL Eur. J. Biochem. 230:752-759(1995).  
CC -!- FUNCTION: THIS GIANT MUSCLE PROTEIN MAY BE INVOLVED IN MUSCLE  
CC ASSEMBLY AND IN MAINTAINING THE STRUCTURAL INTEGRITY OF  
CC SARCOMERES. MAY HAVE PROTEIN KINASE ACTIVITY.  
CC -!- ALTERNATIVE PRODUCTS: A NUMBER OF FORMS OF THIS PROTEIN ARE  
CC PRODUCED BY ALTERNATIVE SPLICING WHICH DIFFER IN TISSUE  
CC DISTRIBUTION. DIFFERENT SIZE TRANSCRIPTS MAY ALSO EXIST WITHIN ANY  
CC ONE TISSUE. THE SEQUENCE SHOWN HERE IS THAT OF THE HEART ISOFORM  
CC N2-B.  
CC -!- TISSUE SPECIFICITY: MUSCLE-SPECIFIC.  
CC -!- SIMILARITY: TO THE CATALYTIC DOMAINS OF OTHER SERINE/THREONINE  
CC KINASES.  
CC -!- SIMILARITY: BELONGS TO IMMUNOGLOBULIN SUPERFAMILY. CONTAINS 112  
CC IMMUNOGLOBULIN C2-LIKE DOMAINS AND 132 FIBRONECTIN TYPE III-LIKE  
CC DOMAINS.  
DR EMBL; X64698; CAA45939.1; -;  
DR EMBL; X83270; CAA58243.1; -;  
DR EMBL; X64697; CAA45938.1; -;  
DR EMBL; X90568; CAA62188.1; -;  
DR EMBL; X64699; CAA45940.1; -;  
DR HSSP; P56276; ITLK.  
DR InterPro; IPR000282; Cytok\_receptor\_2.  
DR InterPro; IPR000719; Euk\_pkinase.  
DR InterPro; IPR000577; FGGV\_kin.  
DR InterPro; IPR003962; FNIII\_repeat.  
DR InterPro; IPR003961; FNIII.  
DR InterPro; IPR001092; HLH\_basic.  
DR InterPro; IPR003598; Ig\_C2.  
DR InterPro; IPR003600; Ig\_like.  
DR InterPro; IPR003006; Ig\_MHC.  
DR InterPro; IPR000129; Peptidase\_S24.  
DR InterPro; IPR002016; Peroxidase.  
DR InterPro; IPR002290; Ser\_thr\_pkinase.  
DR InterPro; IPR001245; Tyr\_pkinase.  
DR Pfam; PF00041; fn3; 132.  
DR Pfam; PF00047; ig; 91.  
DR Pfam; PF00069; pkinase; 1.  
DR PRINTS; PR00014; FNTYPEIII.  
DR PRINTS; PR00726; LEXASERPTASE.  
DR ProDom; PD000001; Euk\_pkinase; 1.  
DR SMART; SM00060; FN3; 127.

```

DR SMART; SM00408; IGC2; 23.
DR SMART; SM00410; IG_like; 79.
DR SMART; SM00220; S_TG; 1.
DR PROSITE; PS00533; FCGY_KINASES_1; UNKNOWN_1.
DR PROSITE; PS00038; HELIX_LOOP_HELIX; UNKNOWN_1.
DR PROSITE; PS00290; IG_MHC; UNKNOWN_1.
DR PROSITE; PS00435; PEROXIDASE_1; UNKNOWN_1.
DR PROSITE; PS00011; PROTEIN_KINASE_DOM; 1.
DR PROSITE; PS00109; PROTEIN_KINASE_TYR; 1.
KW Muscle protein; Cytoskeleton; Structural protein; Calmodulin-binding;
KW Serine/threonine-protein kinase; Alternative splicing; Repeat;
KW Immunoglobulin domain; Phosphorylation.
FT DOMAIN 1370 1389
FT 4 X 5 AA TANDEM REPEATS OF R-M-S-P-A.
FT 429 4614
FT 24731 25070
FT 25030 25056
FT 1372 1372
FT 1377 1377
FT 1382 1382
FT 1387 1387
FT 26171 26171
FT 26178 26178
FT 26184 26184
FT 26190 26190
FT 22277 22277
FT 22449 22449
FT 22454 22454
FT 23324 23324
FT 26926 AA; 2993428 MW; DSEECD3254DF5523 CRC64;
SQ SEQUENCE 26926 AA; 2993428 MW; DSEECD3254DF5523 CRC64;

Query Match 34.7%; Score 463; DB 4; Length 26926;
Best Local Similarity 0.7%; Pred. No. 0.00021;
Matches 168; Conservative 43; Mismatches 34; Indels 23731; Gaps 53;

QY 2 WLK----- 4
Db 1543 WLKNSDIIVPHKPIRIEKTGEAALKIDSTVSQDSAWYTATANKAGRDTRCKVNV 1602
QY 5 -----AFYDK----- 9
Db 1603 VFAPEPEPERKLIIPRGTYRAKEIAAPELEPLHRYGQEWEEGLDYDKEKQKPFKKK 1662
QY 10 ----- 9
Db 1663 LFSRLKRFPAHFCRLTPISDPTMVVLEWLDHGKPLEAANRLRMINEFGYCSLDYVAY 1722
QY 10 ----- 9
Db 1723 SRDSGIITCRATNKYTDTHTSATLIVKDEKSLVESQLPEGRKGLQRIEELERMAHEGAL 1782
QY 10 ----- 9
Db 1783 TGVTTDQKEKQPDIVLYPEPVRVLEGTARFCRCVTGYPPQKYNWNLGQLIRKSRFR 1842
QY 10 -----VAE----- 12
Db 1843 VRYDGIHLDIVDCKSYDTGEVKVTAENPEGVIEHKVKLEIQQRDEFRSLRAPPERPE 1902
QY 13 -----KLKEA----- 17
Db 1903 FHVHEPGKQFEVQKVRPVDTTETKEVVKLRAERITHEKVPESSELSRKFRRTEEG 1962
QY 18 ----- 17
Db 1963 YYEAITAVELSKRKDESYELLRTKTDLLHWTKELTTEEKALAEGBKITPTFKPKD 2022
QY 18 ----- 17
Db 2023 IELSPSMEAPKIFERIQSVTQGSDAHFRVRVVGKPDCECEWKNGVKIERSDRIYWW 2082
QY 18 ----- 17
Db 2083 PEDNVCELVRDVTADSASINVKAINIAGETSSHAFLVQAKQLITFTQELQDVVAK 2142

```

```
QY 25 ----- 24
Db 3223 KISWKEEQLSTGFKCFLHDQEVTLLEAFPEDAAYVCEAKNDYGVATTSASLSV 3282
QY 25 ----- 24
Db 3283 EPEVVSPOEMPVYPPIAITPLQDVTISGQPARQCRVSGDGLKVSWSKDKKIKPSR 3342
QY 25 ----- 24
Db 3343 FFRMTQFEDYQLEIAEAYPEDEGTTFVANNAGQVSSTANLSLEAPESILHERIEQEI 3402
QY 25 ----- 24
Db 3403 EMEMKEFSSFLSAEEGLHSAELQLSKINETLELLSESPVYPTKFDSEKGTGPIFIKE 3462
QY 25 ----- 24
Db 3463 VSNADISMGDVATLSVTIGIPKPIQWPNFNGVLLTPSADYKFVFDGDHSLIILFTKLE 3522
QY 25 ----- 27
Db 3523 DEGEYTCMASNDYKGTICSAYLKINSKGEGHKDTETESAVAKSLEKLGPGCPHFLKELK 3582
QY 28 ----- 27
Db 3583 PIRCAOGLPAIFEYTVVGEPAFTVTFWFKENKQLCTSVIYITIHNPNNGSGTFIVNDPQRED 3642
QY 28 -----GPSV--- 39
Db 3643 SGLYICKAENMLGESTCAABELLVLEDTMDTPCKRAKSTPEAPEDFPQTPLAGPAVEAL 3702
QY 40 ----- 39
Db 3703 DSEQETATFVKDILKAALITEENQOLSYEHIAKANELSSQLPLGAQELQSILOQDKLTP 3762
QY 40 -----FL--- 41
Db 3763 ESTREFLCINGSIHFOPLKEPSNLQIQVQSOKTFSKEGILMPPEPETQAVLSDEKIF 3822
QY 42 ----- 41
Db 3823 PSAMSEIQINSLAVEPLKTLLAEPEGNYPOSSIEPPMHSYLTSVABEVLSEKTEKTSDTN 3882
QY 42 ----- 41
Db 3883 REORVTLQOEQASALILSOSLAEGHVESLQSPDVMISQVNYEPLVPSEHSCHEGGKILI 3942
QY 42 -----FP--- 43
Db 3943 ESANPLENAGODSAVRIEKGSLRFPPLALEEKQVLLKEEHSNDNVMPDQIIESKREPVA 4002
QY 44 ----- 43
Db 4003 IKKQVEQGRDLLSKESLLSGIPEEQRNLKIQICRALQAAVASEQPLGFSEWLRNIEKV 4062
QY 44 ----- 43
Db 4063 EVEAVNITQEPRHIMCMYLVTSAKSVTEEVTIIEDVDPMANLKMELRDALCAIYYEEI 4122
QY 44 ----- 43
Db 4123 DILTAEGPRIQOAKTSLQEMDSFGSQKVEPITEPEVESKYLISTEEVSYNVQSRVK 4182
QY 44 ----- 43
Db 4183 YLDATPVTKGVAVSDEKQDES LKPSKEESESSEGTTEVATVKIQEAGGLIKEDG 4242
QY 44 ----- 43
Db 4243 PMIHTPLVDTVSEGDIVHLTTSITNAKEVNWYFENKLVPSDEKFKCLQDQNTYTLVIDK 4302
QY 44 ----- 43
```

```
Db 4303 VNTEDHQGEYVCEALNDSGKTATSAKLTVVKRAAPVIKRIEPLVALGHLAKFTCEIQS 4362
QY 44 ----- 43
Db 4363 APNVRQWFQKAGRIEYEDSKSTRSSKYISSLEILRTQVVVDGGEYCKASNEYGSVCTA 4422
QY 44 -----PKPKDTLMI--- 52
Db 4423 TLTVTVPGEKKVKRLLPERKPEKEEVVLKSVLRKRPEEEEPKVEPKKLEKVKKPAVPE 4482
QY 53 ----- 52
Db 4483 PPPPKPVEEVEVTVTKRKRKIPEPTKVPEIKPAIPALPAPEPKPKPEAEVKTIKPPVPEP 4542
QY 53 -----SRTP--- 56
Db 4543 EPTPIAAPVTVPVVGKAAEAKAPEAKPKGPIKGVPKKTPSPPIEAERKRLRPGSGEK 4602
QY 57 ----- 56
Db 4603 PPDEAPFTYOLKAVPLKFVKYKIDILITESEFVGSSAIFECLYSPSTAITTNMKDGSNIR 4662
QY 57 -----EYTCV--- 61
Db 4663 ESPKHRFIADGDKRKLHIIDVQLSDAGEYTCVLRNLGNKEKTSTAKLVVEELPVRVKTL 4722
QY 62 ----- 61
Db 4723 EEVTVVGQPLYLSCELNKERDVMVRKDGKIVVEKFRIVPGVIGLMRALTIINDADTDA 4782
QY 62 ----- 61
Db 4783 GTVTVTVENANNLECSKVVEVIRDWLVKPIRDQHVKPKGTAFACDIATKPTNIKWF 4842
QY 62 ----- 61
Db 4843 KGYDEIPAEPNDKTEILROGNHLYLKIKNAMPEDIAEYAVEIEGKRYPAKLTJGEREVEL 4902
QY 62 ----- 61
Db 4903 LKPIEDVTIYKEKESAFDAEISEADIPGOWKLGELLRPSPTCEIKAEKGKRFLLHKVK 4962
QY 62 ----- 61
Db 4963 LDQAGEVLYQALNAITAITLTVKEIELDFAVPLKDVTPERRQARPECVLTREANVWSK 5022
QY 62 ----- 61
Db 5023 GPDIIKSSDKFDIIADGKKHILVINDSQDFDEGYTAEVEGKKT SARLFTVGTIRLAFMSP 5082
QY 62 -----VVDVSHE--- 68
Db 5083 LEDQTVKEGTATFVGCLSHEKMHVVMFKNDAKLHTSRVTLISSEGKTHKLEMKEVTLDD 5142
QY 69 ----- 68
Db 5143 ISQIKAQVKELSSSTAQLKVLADPYFTVKLHDKTAVEKDEITLKCEVSKDVPVWPKDGE 5202
QY 69 ----- 68
Db 5203 EIVPSPKYSIKAGLRLRIKKADLKDGKEYVDCGTDKTKANVTVEARLIEVEKPLYG 5262
QY 69 ----- 68
Db 5263 VEVFVGETAHEIELSEPDVHGOWKLGQPLTASPDCEIIEDGKKHILILHNCQLGMTGE 5322
QY 69 ----- 68
Db 5323 VSTQAAANAANAANKVKELPLIFITPLSDVKVFEKDEAKFECEVSREPKTFRMLKGTQEI 5382
QY 69 ----- 68
```

Db 5383 TGDREFELIKDGTKHSMVIKSAFADEAKYMFEDKHTSGKLIIEGIRLKFLTPLKDVT 5442  
QY 69 ----- 68  
Db 5443 AKEKESAVTVVELSHDNIRVKFNKNDORLHTRSVSMQDEGKHTSITFKDLSIDTQSIR 5502  
QY 69 -----DP----- 70  
Db 5503 VEAMGMSSEAKLTVLEGGDPYFTGKLQDYGVEKDEVILQCEISKADAPVKWFKDGKEIKP 5562  
QY 71 ----- 70  
Db 5563 SKNAVIKTDGKKRMLILKALKASDYGQYTCDCGTDKTSGLKDIEDREIKLVRPLHSVEYM 5622  
QY 71 ----- 70  
Db 5623 ETETARFETEISEDDIHANWKLGALLQTPDCEIKEBKGIIHSLVLHNCRLDQTCGVDPQ 5682  
QY 71 ----- 70  
Db 5683 AANVKSSAHLRVKPRVIGLLRPLKDVTVTGETATDFCELSYEDIPVEWYLGKKLEPSD 5742  
QY 71 ----- 70  
Db 5743 KVVPRSEGVHTLRLRVKLEDAAGEVQLTAKDFKTHANLKVKEPPVEFTKPLEDQTVBEG 5802  
QY 71 ----- 70  
Db 5803 ATAVLECEVSRENKVKWFKNGTEILKSKYIEIVADGRVKLVIHDCPTEDIKTYTCDAK 5862  
QY 71 ----- 70  
Db 5863 DFKTSCNLVPPHVEFLRPLTDLQVREKEMARFECELSRENKVKWFKDGAIEIKKGRKY 5922  
QY 71 ----- 70  
Db 5923 DIISKAVRILVINKCLLDDEAEYSCEVTRTARTSGMLTVLEBEAVFTKLANIEVSETDT 5982  
QY 71 ----- 70  
Db 5983 IKLVCEVSKPGAELVIWYKGDDEIIETGRVEIITTEGRKRLIVIONAHLEDAGNYNCRLPSS 6042  
QY 71 ----- 70  
Db 6043 RTDGKVKVHELAAEFISKPNLEILEGEKAEPVCSISKESFPVQWKRODKTLESKDYDV 6102  
QY 71 -----EYKF 74  
Db 6103 IADGKKRVLVVKDATLQDMGTVVMVGAARAAAHLTVEKLRIVVPLKDRVKEQOEYVF 6162  
QY 75 N----- 75  
Db 6163 NCEVNTGAKAKAFRNEEAIFDSSKYIILQKDLVTLRLIRDAHLDQANYVSLTNHRGE 6222  
QY 76 ----- 75  
Db 6223 NVKSAANLIVEBEDLRIVEPLKDIETMEKKSVTFWCKVNRNLNVTWKTKNGEEVFPDNRY 6282  
QY 76 ----- 75  
Db 6283 SYRVDKXKHLMTIKDCGFPDEGEYIVTAGQDKSVAELLIIIEAPTVEFVEHLEDQTVTEFDD 6342  
QY 76 ----- 75  
Db 6343 AVFSCQLSREKANVXWYRNGREIKEGKKYKPEKDGSIHRLIIKDCRLDDECEYACGVEDR 6402  
QY 76 ----- 75  
Db 6403 KSRARLFVEEIPVEIIRPPQDILEAPGADVVFVLAELNKNKVEQVLRNRMVWVGDKHQH 6462  
QY 76 ----- 75  
Db 6463 MSEGKIHRLQICDIKPRQGEYRFTAKDKEARAKLELAAAPKIKTADQDLVVDVGKPLTM 6522

QY 76 ----- 75  
Db 6523 VVPYDAYPKAEAEWFKENEPLSTKTIDTAEQTSFRILEAKKGDKGRKIVLONKHGKAE 6582  
QY 76 ----- 75  
Db 6583 GFINKLVIDVPGVPRNLEVTFDFGEVSLAWBEPLTDGSGKIIGYVVERRDIKRTKWVLA 6642  
QY 76 ----- 75  
Db 6643 TDRASCEFTVTGLQKGGVEYLFVRSARNRVGTGEPVETDNPVEARSKYDVPGLPNTVI 6702  
QY 76 ----- 75  
Db 6703 TDVNRFGVSLTWEPPEYDGGAEITNYVIELRDKTSIRNDTAMTVRAEDLSATVTDVVEGO 6762  
QY 76 ----- 75  
Db 6763 EYSFRVRAQNRIGVGKPSAATPFVKVADPIERPSPPNLTSSDQTSQSSVQLKWEPLKDG 6822  
QY 76 ----- 75  
Db 6823 GSPILGYIIERCEEGKDNWIRCNMKLVPETYKVTGLEKGNKYLYRVSANENKAGVSDPSE 6882  
QY 76 ----- 75  
Db 6883 ILGPLTADDAFVEPTMDLSAFKDGLEVIVPNPITILVPSTGYPRPTATWCFGDKVLETGD 6942  
QY 76 ----- 75  
Db 6943 RVKMKTL SAYAELVISPSERSDKGIYTLKLENRVKTIISGEIDVNVVIARPSAPKELKFGDI 7002  
QY 76 ----- 75  
Db 7003 TKDSVHLTWEPDDGGSPLTGYVVEKREVSRTKTKVMDVFVTDLEFVDPDLVQGEKYLEF 7062  
QY 76 ----- 75  
Db 7063 KVCARNKCGEPAYVDEPVNKMSTPATVPDPENVKWRDRTANSIFLTWDPKPDGSGRI 7122  
QY 76 ----- 75  
Db 7123 KGYIVERCPGRGSKWVACGEPVAETKMEVTGLEEGKWYAVRVTLNROGAKSPSRPTBEI 7182  
QY 76 ----- 75  
Db 7183 QAVDTQEAPEIFLDVKLLAGLTVKAGTKIELPATVTKGPEPKITWTKADMILKQDKRITI 7242  
QY 76 ----- 75  
Db 7243 ENVPKSTVTIVDSKRSDTGTIIIEAVNVCGRATAVVEVNVLDKPGPAAFDITDVTNES 7302  
QY 76 ----- 75  
Db 7303 CLLTWNPPRDDGGSKITNYVVERRATDSEVWHKLSSTVKOTNFKATKLIPNKEYIFRVA 7362  
QY 76 ----- 75  
Db 7363 ENMYGAGEPVQASPTAKYQFDPGPPTRLEPSDITKDAVTLTWCEPDGGSPIYGVW 7422  
QY 76 ----- 75  
Db 7423 ERLDPDTDKWYRCNKMPVKDITYRVKGLTNKKYRFRVLAENLAGPGKPSKSTBELLIKD 7482  
QY 76 ----- 75  
Db 7483 PIDPPWPBGKPTVKDVGKTSVRLNWTKEPHDGGAKIESYVIELMLKTGTDEWVRVAEGVPT 7542  
QY 76 ----- 75  
Db 7543 TQHLPLGLMEQGEYSFRVRAVNKAGESEPSDPVLCREKLYPPSPRWRLEVINITNKT 7602

QY	76	-----	75
Db	7603	ADLKWTPEKGGSPITNIVEKRDYRRKGQWTVDTTKCTVTPLEGSLYVFRVAA	7662
QY	76	-----	75
Db	7663	ENAIQSDYTEIEDSVLAKDTFTTGPYPYALAVVDVTKRHVDLKBPPKNDGGRPIQRYV	7722
QY	76	-----	75
Db	7723	IEKKERLGRWVRKAGTAGDCNFRVTDVIEGTEVOFQVRAENEAGVGHPSPTLSIE	7782
QY	76	-----	75
Db	7783	DPTSPSPPLDLHVTDAGRKHIAIAWKPKPKNGGSPILGYHVEMCPVGTEKWRVNSRPI	7842
QY	76	-----	75
Db	7843	KDLKFRVEGVDPKEYVLRVRAVNAIGVSEPSEISENVVAKDPDCKPTIDLETHDIIVI	7902
QY	76	-----WYVDG-----	80
Db	7903	EGEKLSIPVFPRAVPVTVSMHKDGKEVKASDRLTMKNDHISAHLEVPKSVRADAGIYTI	7962
QY	81	-----	80
Db	7963	TLENKLSATASINVKVIGLPGPKDIKASDITKSSCKLTWEPPEFDGGTPILHYVLERR	8022
QY	81	-----	80
Db	8023	EAGRTYIPVMGENKLSWTVDLIPNGEYFFRVKAVKVGGEYIELKNPVIAQDPKQP	8082
QY	81	-----VEVNA-----	86
Db	8083	PDPPVDVEVHNPTAEAMTITWKPPLYDGGSKIMGYIIIEKIAKEERWKRCHLVPILTY	8142
QY	87	-----	86
Db	8143	TAKGLEEGKEYQFVRRAENAGISEPSRATPPTKAVDPIDAPKVILRTSLEVRGDEIAL	8202
QY	87	-----	86
Db	8203	DASISGSPYTIWIKDENVIVPEEIKKRAAPLVRRRKGEVQEEPPVLPTQRLSIDNS	8262
QY	87	-----	86
Db	8263	KGESQLVRDSLRPDHGLYMIKVENHDGIAKAPCTVSVLDTPGPPINFVFEDIRKTSVL	8322
QY	87	-----KTKP-----	90
Db	8323	CKWEPPLDDGGSEIINVTLEKDKTKPDSEWIVVTSTLRHCKYSVTKLIBGKEYLFRVRA	8382
QY	91	-----	90
Db	8383	ENRFGPPPCVSKPLVAKDPGPPDAPDKPIVEDVTSNSMLVKWNEPKDNGSPILGYWLE	8442
QY	91	-----	90
Db	8443	KREVNSTHNSRVNKSLLNALKANVDGLLEGLTVFVRVCAENAAAGPKFSPSPDKTAHDP	8502
QY	91	-----	90
Db	8503	ISPPGPPIPRVTDTSSTTIELEWEPFAPNGGGEIVGYFDKQLVGTNKNWRCCTEKMVKVR	8562
QY	91	-----	90
Db	8563	QYTVKEIREGADYKLRVSAVNAAGEGPPGETQPVTVAEPOEPFAVELDVSVKGGIQUIMAG	8622
QY	91	-----	90
Db	8623	KTLRIPAVVTGRPVTKVWTKEGELDKDRVIDNVGTKSELIIKDALRKDHGRYVITAT	8682
QY	91	-----	90

Db	8683	NSCGSKFAARVEVFDVPGPVLDLKPVTNKRKMLLNWSDPEDDGGSEITGFIIERKDAK	8742
QY	91	-----	90
Db	8743	MHTWROP IETERSKCDITGLLEGOEYKFRVIAKNKFGCGPPVEIGPILAVDPLGPPTSPE	8802
QY	91	-----	90
Db	8803	RLTYTERQSTITPLDWKEPRSGSGPIQGYIIIEKRHRDKPDFERVNKRCLCPTTSFLVENL	8862
QY	91	-----	90
Db	8863	DEHQMEYFRVKA VNEIGESEPSPLNVV IODDEVPPPTIKRLSVRGDTIKVKAGEPVHIP	8922
QY	91	-----	90
Db	8923	ADVTGLPMPKIEWSKNETVIEKPTDALQITKEEVSSEAKTELSIPKAVREDKGTYYTVA	8982
QY	91	-----	90
Db	8983	SNRLGSFERNVHVVEDRPSPPRNLA VTDIKAESCYLTWDAPLDNGGSEITHYVIDKRDA	9042
QY	91	-----	90
Db	9043	SRKKAWEETNTAVEKRYGIWKLIPNGOYEFVRVAVNKYGISDECKSKVVIQDPYRLP	9102
QY	91	-----	90
Db	9103	GPPGPKVLARTKGMVLSWTPPLDNGGSPITGYWLEKREEGSPYMSRVSRAPITKVGLK	9162
QY	91	-----	90
Db	9163	GVFENVRLLEGVYOFRAMAINAAGIPSPSDPEVAGDPIFPPOGPPSCPEVKDKTKS	9222
QY	91	-----	90
Db	9223	SISLGHKPPAKDGGSPILGYIVEMQEBGTTDWKRVNEPDKLITTCCEVVPNLKELRYRF	9282
QY	91	-----	90
Db	9283	RVKAVNAGESEPSDTTGEIPATDIOEPEVFIDIGAQCLVCKAGSQIRIPAVIKGRPT	9342
QY	91	-----	90
Db	9343	PKSSEFPDGAKKAMKGVHDIPEDAOLEAENS SVIIIECKRSHTGKYSITAKNKAGQ	9402
QY	91	-----	90
Db	9403	KTANCRVKVMDVPKPKDLKAVSDITRGSRLSNKMPDDDDGGDRIKGVVIEKRTIDGKAWT	9462
QY	91	-----REEQY-----	95
Db	9463	KVNPDCGSTTFVVPDLLSEQOYFFRVRAENRFGIGPPVETIQRTARDPIYPPDPIKLK	9522
QY	96	-----	95
Db	9523	IGLITKNTVHLSWPKPKNDGSGSPVTHYIVECLAWDPTGTTKEAWRQCNKRQVBELOQTFVE	9582
QY	96	-----	95
Db	9583	DLVEGGEYEFVRKAVNAAGVSKPSATVGPCDCORPMDPPSIDLKEFMEVEEGTNVIVAK	9642
QY	96	-----	95
Db	9643	IKGVPPFTLTFWKAPPKPDNKEPVLVYTHVNKLVDVDTCTLVIPQSRSDTGLYTITAV	9702
QY	96	-----	95
Db	9703	NNLGASKEMRLNVLRGPPGVPKIFESVSADQMTLSWTFPPKDDGGSKITNVVIEKREA	9762
QY	96	-----	95

Db 9763 NRKTVHVSSEKECTYTPKLLLEGHYVFRIMAQNKYIGIBPLDSEPETARNLFSVPGA 9822  
QY 96 ----- 95  
Db 9823 PDKPTVSSYRNSMTVNWEEPEYDGGSPYGYWLEKMDTTSKRWRVNRDPIKAMTLGYS 9882  
QY 96 ----- 95  
Db 9883 YKVTGLIEGSDYQFVYAINAAGVPASLPSPATARDPIAPPGPPKVTWTKSSADL 9942  
QY 96 ----- 95  
Db 9943 EWSPPPLKOGSKVTGYVIEYKEEGKEWKGDKKEVGRGKLVVTGLKEGAFYKFRVSAVN 10002  
QY 96 ----- 95  
Db 10003 IAGIGEGETVDIEMKORLVSPDLQDASVRDRIVVHAGVIRIAYVSGKPPPTVTWN 10062  
QY 96 ----- 95  
Db 10063 MNERTLPQEAETIETTAISSSMVINKQCRSHQGVYSLAKNEAGERKKTIIIVDLDVPGPV 10122  
QY 96 ----- 95  
Db 10123 GTPFLAHLNLTNSCKLTWFSPEDDGSPITNYVIEKRESRRAWTPVTYVTRQATVQG 10182  
QY 96 ----- 95  
Db 10183 LIQGRAYFPRIAAENSIGMPFVETSEALVIREPITVPERPEDLEVKEVTKNTVTLTWNP 10242  
QY 96 ----- 95  
Db 10243 PKYDGSSEIYNVLESRLIGTEFKHKVTNDNLLSRKTYVKLKEGDTVEYRVSAVNIQO 10302  
QY 96 ----- 95  
Db 10303 GKPSFCTKPTCKDELAPTLHLDFRDLKUTIRVGEAFALTGRYSGKPKVSWFDEADY 10362  
QY 96 ----- NST----- 98  
Db 10363 LEDDRTHIKTTPATLAEKIKAKRSDSGKCYVVENSTGSRKGFQCVNVVDHPGPPGV 10422  
QY 99 ----- 98  
Db 10423 SFDEVTKDYWISWKPPLDGGSKITNYLIEKKEVGDVWMPVTSASAKTTCKVSKLLEG 10482  
QY 99 ----- 98  
Db 10483 KDYIFRIHAENLYGSDPLVSDSMKAKDRFRVPDAPDQPIVTEVTKDSALVTWNKPHDGG 10542  
QY 99 ----- 98  
Db 10543 KPITNYILEKRETSKRWARVTKDPIHPYTKFRVPDLLEGCOYEFPRVSAENEIGIDPSP 10602  
QY 99 ----- 98  
Db 10603 PSKVPFAKDPYAKSPVPVNPPEAIDTTCNSVOLTWQPRHDGSKILGYVIEYQKVGDSEW 10662  
QY 99 ----- YRVVS----- 104  
Db 10663 RRANHTPESCPTKYKVTGLRDGQYKFRVLAVNAAGESPAHVPEPVLVKDRLEPPELI 10722  
QY 105 ----- 104  
Db 10723 LDANMARQHIKVGDTLRLSALIKGPPPKVTWKKEDRDAPTAKRIDVTPVGSKLEIRNA 10782  
QY 105 ----- LTV----- 107  
Db 10783 AHEDGGIYSLVTENPAGSKTVSVKVLVDKPGPRDLEVSEIRKDCSYLTNKEPLDDGGS 10842  
QY 108 ----- 107  
Db 10843 VITNTVVERRDVASQWSPLSATSKKSHFAKHLNEGNOYLFRVAENQYGRGPPVETPK 10902

QY 108 ----- LH----- QDWL-- 113  
Db 10903 PIKALDPLHPPGPKDLHHVDVDKTEVSLVWNKPRDGGSPITGYLVYEYQEGTQDWIKF 10962  
QY 114 ----- 113  
Db 10963 KTVTNLECVVTGLQOGKTYRFRVKAENIVGLGLPDTTPIEEOEKLVPSPVELDKLIBG 11022  
QY 114 ----- 113  
Db 11023 LVVKAGTTVRFPAILIRGVPVPTAKWTTDQSEIKTDEHYTVETDNFSSVLTIKNCLRRDTG 11082  
QY 114 ----- 113  
Db 11083 EYQITVSNAAGSKTVAVHLTVLDVPCPTGPIINILDTVPEHMTISWQPKDDGGSPVINY 11142  
QY 114 ----- 113  
Db 11143 IVEKQDTRKDTWGVVSSGSKTKLKIPLQKGEYVFRVRAENKIGVGPPLDSTPTVAKH 11202  
QY 114 ----- 113  
Db 11203 KFSPPSPCKPVVTDITENAATVSWTLPKSDGGSPIGTGYMERREVTGKWRVNKPTIAD 11262  
QY 114 ----- 113  
Db 11263 LKFRVTGLYEGNTYBFRVFAENLAGLSKSPSSDPIKACRPKIPGPPINPKLKDKSRT 11322  
QY 114 ----- 113  
Db 11323 ADLVWTKPLSDGSPILGVVVECOQPGTAQNNRINKDELIOCAFRVPGLEGNEVFRFI 11382  
QY 114 ----- 113  
Db 11383 KAANIVGEPEPRELAESVIAKDIHLHPPEVELDVTCRDVTIRVGTOTIRLARVKGRPEPD 11442  
QY 114 ----- 113  
Db 11443 ITWTKGKVLVREKRVLDLIQDLPRVELQIKEAVRADHGKYIISAKNSSCHAOGSAIVNVL 11502  
QY 114 ----- 113  
Db 11503 DRPGCQNLKVTNVTKENCITISWENPLDNGGSEITNFIVEYKRNKOGKWSIVASDVTKRL 11562  
QY 114 ----- 113  
Db 11563 IKANLLANNEYFRVCAENKVGCGPTIETKTPILAINPIDRGPENLHIAKGTFFVYL 11622  
QY 114 ----- 113  
Db 11623 KWRRPDYDGGSPNLSYHVERRLKGSDDHVHKGSIKETHYVNDRCVENQIYEFVRQTKN 11682  
QY 114 ----- 113  
Db 11683 EGGSDWVKTEBVVVKEDLQKPVLDLKLGLVLTAKGDTIRLEAGVRGKPPPEVAVTKDK 11742  
QY 114 ----- 113  
Db 11743 DATDLTRSPPRKIDTRADSSKFSLTAKRSDGKYYVVTATNTAGSFVAYATVNVLDKPGP 11802  
QY 114 ----- 113  
Db 11803 VRNLKIVDVSSDRCTVCWDPPEDDGGCEIQNVILEKCETKRWVNSTYSATVLTPTGTVTR 11862  
QY 114 ----- 113  
Db 11863 LIEGNEYIFRVAENKIGTGPPTESKPIAKTKYDKGRDPPPEVTKVSKSEMTVWVNNP 11922  
QY 114 ----- 113  
Db 11923 EYDGGKSTGYFLEKKKXKSTRVVPVNSAIPERRMKVQNLPLPDHEYQFRVAENEIGIG 11982

QY 114 ----- 113  
Db 11983 EPSLSPRVVAKDIEPPGPPTNFRVVDTKHSITLWGKXPYDGGAPIIGYVEMRPKI 12042  
QY 114 ----- 113  
Db 12043 ADASPDGKRCNAALVRKEFTVTSLDENQEYEFRCVCAQNOVGIGRPAELKEAIKPK 12102  
QY 114 ----- 113  
Db 12103 ILEPPEIDLASMRKLIVIRAGCPIRLFAIVRGRPAKVTWRKVGDINNVVRKGOVDLVT 12162  
QY 114 ----- 113  
Db 12163 MAFIVIPNSTRDDSGKYSLTLVNPAGEKAVFVNVRLDTPGPVSDLKVSDVTKTSCHVSW 12222  
QY 114 ----- 113  
Db 12223 APPENDGGSQVTHYIVEKREARKTWSTVTPVKKTSFHVTLNLPNGEYFRVTAUNEYG 12282  
QY 114 ----- 113  
Db 12283 PGVPTDVKPVLASDPLSEDPKPKLEATEMTKNSATLAWLPLRGGAKIDGYIISYRE 12342  
QY 114 ----- 113  
Db 12343 EEQPADRWTEYSVVKDLVLVVTGLKEGKKYKFRVAARNAVGVSLPREAGVYEAKQLLP 12402  
QY 114 ----- 113  
Db 12403 PKILMPQITIKAGKKLRIEAVYGRPHPTCKWKKGDEVVTSVSHLAVHKADSSILIK 12462  
QY 114 ----- 113  
Db 12463 DVTRKDSGYSLTAENSSGTDQIKIVVMDAPGPPQPPEDISDIDADACSLSWHIPLE 12522  
QY 114 ----- 113  
Db 12523 GGSNITNYIVEKCDVSRGDVNTALASVTKTSCRVGKLIIPQGEYIFRVAENRFGISEPLT 12582  
QY 114 ----- 113  
Db 12583 SPRMVAQFPFGVPSEPKNARVTKVNDKCFVAMDRPDSGGSPIIGYLIBRBERKNSLLW 12642  
QY 114 ----- 113  
Db 12643 KANDTLVRSTPCAGLVEGLEYSFRYIALNKAGSSPPSKPTXYVTARMPVDPGKPEVI 12702  
QY 114 ----- 113  
Db 12703 DVTKSTVSLIWARPKHDGSKIIGYFVEACKLPDGKWRNCNTAPHQIPOEYATGLEEK 12762  
QY 114 ----- 113  
Db 12763 AQYQFRAIARTAVNISPPSEDPVTTLAENVPRIIDLVSAMKSLTAVKAGTNVCLDATV 12822  
QY 114 ----- 113  
Db 12823 FGKPMPTVSWKKDGTLLKPAEGIKMAMQORNLCITLFLFSVNRKDSGDYTTIATENSSGSKSA 12882  
QY 114 ----- 113  
Db 12883 TIKLVLDKPPASVAKINKMYSRAMLSWEPPLDGGSEITNYIVDKRETSRPNWAQVS 12942  
QY 114 ----- 113  
Db 12943 ATVPITSCSVEKTEGHEYQFRICAENKYGVDPVFTEPALAKNPYDPPGCDPPVISNI 13002  
QY 114 ----- 113  
Db 13003 TKDHMTVSWRPPADGGSPITGYLLEKRETQAVNWKVNRKPIIERTLKATGLOEGTEYE 13062  
QY 114 ----- 113

Db 13063 FRYTAINKAGPGKPSDASKAAYARDPOYPPAPPAPFKVYDTRSSVSLSMGKPAYDGGSP 13122  
QY 114 ----- 113  
Db 13123 IIGYLVEVKRADSDNVVRCNLQNLOKTRFEVTGLMEDTOYQFRYAVANKIGYSDPSDVP 13182  
QY 114 ----- 113  
Db 13183 DKHYPKDILIPPEGEHDADLRKTLILRAGVTMRLYVPVKRPPPKITWSPKNWLRDRIG 13242  
QY 114 ----- 113  
Db 13243 LDIKSTDFDLRCENVNKYDAGKYILTLNENSCGKKEYTIWVKVLDTPGPPINVTVKEIS 13302  
QY 114 ----- 113  
Db 13303 KDSAYVTWEPPIIDGSPIIINVVQKRDAERKSMSTVTTECKSTFRVPNLEEGKSYFFR 13362  
QY 114 ----- 113  
Db 13363 VFAENEYIGIDPGETRDVAKASQTPGPVVDLKVRSVSKSSSIGWKPHSDGGSRIIGYV 13422  
QY 114 ----- 118  
Db 13423 VDFLTEENKQWVMKSLSLQYSAKDLTEGKEYTFRVSAENENGEGBTPSEITVWARDVVA 13482  
QY 119 ----- 118  
Db 13483 PDLDLKGLPDLCLAKENSFRLKPIKGRPAPSVSMKKGEDPLATDRVSESSAVNTT 13542  
QY 119 ----- 118  
Db 13543 LIVVDCOKSAGKVTITLKNVAGTKECTISIKVVGKPGIPTGPIKFDEVTAEAMTLKWP 13602  
QY 119 ----- 118  
Db 13603 PKDDGGSEITNYILEKRDSVNNKWTCASAVQKTTFRVTRLHEGMEYTFRVSAENKYGVG 13662  
QY 119 ----- 118  
Db 13663 EGLSEPIVARHPDPVDPAPPNNIVDRHDSVSLTWTDPKKTGSPITGVHLEFKERNS 13722  
QY 119 ----- 118  
Db 13723 LLWKRANKTPIRMDFKVTGLTEGLEYEFRVMAINLAGVKPSLPSEPVVALDIDPPGK 13782  
QY 119 ----- 118  
Db 13783 PEVINITRNSVTLIWTPEKYDGGHKLTYIVEKRDLPFSKMMKANHVNPCEAFTVTDLV 13842  
QY 119 ----- 118  
Db 13843 EGGKYEFRIRAKNTAGIASPSESTETIICKDEYEAPTIVLDPTIKDGLTIKAGDTIVLN 13902  
QY 119 ----- 118  
Db 13903 AISILGRLPLPKSSWSKAGKDIRSDITQITSTPTSSMLTIKYATRKDAGEYITITATNPF 13962  
QY 119 ----- 118  
Db 13963 TKVEHVKTVDVPPGPPVEISNVSAEKATLTWTPPLEDGGSPIKSYILEKRETSRLW 14022  
QY 119 ----- 118  
Db 14023 TVWSEDIQSCRHVATKLIQGNIEYIFRVSAVNHYGEPVQSEPVKMWDRFGPPGPEKPE 14082  
QY 119 ----- 124  
Db 14083 VSNVTKNTATVSWRKPVDGSEITGYHVERREKSLRWRAITPVSDDLCKVTLGQEG 14142  
QY 125 ----- 124

Db 14143 STYEFVRSAENRAGIGPPSEASDSVLMKDAAYPPGPPSPHVTDTTKKSASLAWGKPHYD 14202  
QY 125 ----- 124  
Db 14203 GGLEITGVVVEHQVGDEAWIKDTGTGALRITQFVVVPLDQTKYKYNFRISAINDAGVGE 14262  
QY 125 ----- KALPAP----- 130  
Db 14263 AVIPDVEIVEREMAPDFELDAELRRTLVRAGLSIRIFVPIKGRPAPEVTWTNDINLKN 14322  
QY 131 ----- 130  
Db 14323 RANIENTESPTLLIIEPCNRYDTGKFWMTIENPAGKSGFVNVVRVLDTPGPVLNLRPTDI 14382  
QY 131 ----- IEK----- 133  
Db 14383 TKDSVTLHNDLPLJDGSRITNIVEKREATRKSYSATTCKHKCYKVTGLSEGCEYFF 14442  
QY 134 ----- 133  
Db 14443 RVMAENEYIGIPEPTETEPVKASEAPSPDLSNIMDITKSTVSLAWPKPKHDGGSKITGY 14502  
QY 134 ----- 133  
Db 14503 VIEAQRKSGDQWTHITTVKGLECVVRNLTEGEYTFQVMVNSAGRSAPRESRPVIVKEQ 14562  
QY 134 ----- 133  
Db 14563 TMLPELDLRGIYQKLVIAKAGDNKVEIPVLGRPKPTVTKKGQDQILKQTORVNFETTAT 14622  
QY 134 ----- 133  
Db 14623 STILNINECVRSDBGPYPLTARNIVEGVGVITIQVHDIPGPTGPKFDEVSSDEVTS 14682  
QY 134 ----- 133  
Db 14683 WDPENDGGVPISNYVEMRQDSTTWVELATTVIRTTYKATRLTTGLEYQFVKQAQRY 14742  
QY 134 ----- 133  
Db 14743 GVGPGITSAWIVANYPFKVPGPPTQVAVTKDSMTISWHEPLSDGSGPILGYHVERKE 14802  
QY 134 ----- TISKA----- 138  
Db 14803 RRGILMQTVSKALVPGNIFKSSGLTDGIAEFVRVAENMAGSKSPSEPMALDPIDP 14862  
QY 139 ----- 138  
Db 14863 PGKPVPLNTRHTVTLKWAQPEYTGFKITSYIVEKRDLPNGRWLKANFSNILENEFTVS 14922  
QY 139 ----- 138  
Db 14923 GLTEDAAEFVRVIAKNAAGAIAPPSEPSDAITCRDDVEAPKIKVDVKFDTVLKAGEAF 14982  
QY 139 ----- 138  
Db 14983 RLEADVSGRPPTMEWSKDGKEGTAKLEIKIADFSTNLVKNKSTRDSGAYTLTATNP 15042  
QY 139 ----- 138  
Db 15043 GGFAKHIFNVKVLDRPGPEGPLAVTEVTSEKCVLSWFPPDDGGAKIDHYIVQKRETSR 15102  
QY 139 ----- 138  
Db 15103 LAWNIVASEVQVTKLVTKLLKGNEYIFRVMVKNYGVGEPLSEPLAVNPYPPDPKP 15162  
QY 139 ----- 138  
Db 15163 NPEVTITKDSMVWCWGHDPDGGSEIINYIVERRDKAGQWIKCNKKTLDLRYKVSGL 15222  
QY 139 ----- 138  
Db 15223 TEGHEVEFRIMAENAGISAPSTSPFFYKACDTVFKPGPPGNPRVLDTSRISISIAWNKP 15282

QY 139 ----- 138  
Db 15283 IYDGGSEITGVYVEIALPEDEMQIYTPPAGLAKATSYTITGLTENQEKIRIYAMNSEGL 15342  
QY 139 ----- 138  
Db 15343 GEPALVPGTPKAEDRMLPPEIELDADLRKVVTIRACCTLRFLVPIKGRPDPEVKWARDHG 15402  
QY 139 ----- 138  
Db 15403 ESLDKASIESASSYTLIIVGNVNRFDGKYILTVENSNGSKSAFVNVVRVLDTPGPQDLK 15462  
QY 139 ----- 138  
Db 15463 VKEVTKTSVTLWDPPLLDGGSKIKNYIVEKRESTRKAYSTVATNCHKTSWKVDQLQEGC 15522  
QY 139 ----- 138  
Db 15523 SYFVRVLAENEYIGLPAETAESVKASERPLPPGKITLMDVTRNSVLSWEPKHDGSR 15582  
QY 139 ----- 138  
Db 15583 ILGYIVEMQTKGSKDWATCATVKVTEATITGLIQGEEYSFRVSAQNEKIGISDPQLSVPV 15642  
QY 139 ----- 138  
Db 15643 IAKDLVIPPAFKLLFNTFTVLAGEDLKVDPVFFIGRPTPAVTHWKDNVPLKQTRVNAEST 15702  
QY 139 ----- 138  
Db 15703 ENNSLLTIKDACREDVGHVYVWLKLTNSAGEAIELNVILDKPGPTGPKVMDDEVTAADSI 15762  
QY 139 ----- 138  
Db 15763 LSWGPPKYDGGSSINNYIVEKRDSTTTTWOIVSATVARTTIKACRLTKCEYQFRIAAEN 15822  
QY 139 ----- 138  
Db 15823 RYCKSTYLNSEPTVAQYPFKVPGPCTPVVTLSSRDSMEVQWNEPISDGGSRVIGYHLER 15882  
QY 139 ----- 138  
Db 15883 KERNSILWVKNLKTPIQTKFKTTGLEGEVEYEFVSAENIVGIGKPKVSECYVARDPC 15942  
QY 139 ----- 138  
Db 15943 DPPGRPEAIIVTRNSVTLOWKKPTVDGGSKITGYIVEKKELPEGRMMAKASFTNIIDTHFE 16002  
QY 139 ----- 138  
Db 16003 VTGLVEDHRYEFVRVIARNAAGVFPSESTGALTARDEVDPPIRISMDPKYKDTIVVHAGE 16062  
QY 139 ----- 138  
Db 16063 SPKVDADIYKPIPTIOWIKGQDELSNTARLEIKSTDFATSLSVKDAVRVDSGNYILKAK 16122  
QY 139 ----- 138  
Db 16123 NVAGERSVTNVKVLDRPGPPEGPVVISQVTAEBCKTLAWKPLQDGGSDIINIIVERRET 16182  
QY 139 ----- 138  
Db 16183 SRLVWTVADVQTLSCVKTKLLEGNEYTFRIMAVNKYGVGEPLSEPPVAKNPFVVPDA 16242  
QY 139 ----- 138  
Db 16243 PKAPEVTTVTKDSMIWVWERPASDGGSEITLGYVLEKRDKEGIRWTRCHKRLIGELRLV 16302  
QY 139 ----- 138  
Db 16303 GLIENHDIYEFVRVSAENAGLSEPPSPAYQKACDPIYKPPNNPKVIDITRSSVFLSWS 16362



QY 139 ----- 138  
Db 16363 KPIYDGGCEIOGYIIVEKCDVNVGEWTCTPTTGINKNTNIEVEKLEHEYNFRICAINKA 16422  
QY 139 ----- 138  
Db 16423 GVGEHADVPGPPIIVEEKLAPDIDLLELRKIIINIRAGGSLRLEVPVPIKNGRPTPEVKWGV 16482  
QY 139 ----- 138  
Db 16483 DGEIRDAIIVDTSSFTSLVDNVRDYSKYITLTLENSGTSKSAFVTVRVLDTSPSPVN 16542  
QY 139 ----- 138  
Db 16543 LKVTEITKDSYSITWEPPLLDGGKIKNY IVEKREATRKSYAAVVTNCHKNKSKIDQLQE 16602  
QY 139 ----- 138  
Db 16603 GCSYFRVTAENEYIGLPAQTADPIKVAEVPQPPGKITVDVDRNSVLSWTKPEHDGG 16662  
QY 139 ----- 138  
Db 16663 SKIIQYIVEMQAKHSEKSECARVKSLOAVITNLTQGEYLFPRVAVNNEKGRSDPRSLAV 16722  
QY 139 ----- 138  
Db 16723 PIVAKDLVIEPDKVPFSSYVQVQDLMKEVPISGRPKPTITWTKDGLPLKQTRINVT 16782  
QY 139 ----- 138  
Db 16783 DSDLTLTISIKETHKDDGGYGITVANVVGOKTASIEIVTLDKDPKPGVKFDDVSAES 16842  
QY 139 ----- 138  
Db 16843 ITLSWNPPLYTGGCOITNY IVQKRDTTWTWDDVVSATVARTTLKVKLTGTEYQFRIFA 16902  
QY 139 ----- 138  
Db 16903 ENRYGQSFALSDPIVAQYPYKEPGPTPFATAISKDSMVIONHEPVNNGSPVIGVHL 16962  
QY 139 ----- 138  
Db 16963 ERKERNILATWKVNTIIHDTRQAKNLEEGIEYEFVRYAENIVGVGKASKNSECYVARD 17022  
QY 139 ----- 138  
Db 17023 PCDPGPGPEIMVXRNEITLQWTKPVYDGGSMITGY IVEKKRDLDPGRWMAKSFNTNIETQ 17082  
QY 139 ----- 138  
Db 17083 FTVSGLTEDQRYEFRVIKNAAGAIKSPSDSTGPITAKDEVELPRISMDPKFRDTIVVNA 17142  
QY 139 ----- 138  
Db 17143 GETRLEADVHGKPLPTLEWLRGDKIEESARCEIKNTDFKALLIVKDAIRIDGGQYTLR 17202  
QY 139 ----- 138  
Db 17203 ASNVAGSKSPFVNVKVLDRPGPPEGPVQVGTSEKCSLTWSPPLQDGGSDISHYVVEKR 17262  
QY 139 ----- 138  
Db 17263 ETSRLATVAVSEVVTNSLKVTKLLEGNEYVFRIMAVNKYGVGELESAPVIMKNPFVLP 17322  
QY 139 ----- 138  
Db 17323 GPPKSLEVTNIAKDSMTVCWNRPSDGGSEIIGYIIVEKRDRSGIRWIKCNKRRTDLRLR 17382  
QY 139 ----- 138  
Db 17383 VTGLTEDEYEFVRSNAAGVGPSPATVYKACDPVKPGPPTNAHIVDTTKNSITLA 17442  
QY 139 ----- 138

Db 17443 WGKPIYDGGSEILGVVVEICKADEEEMQIVTPQTGLRVTRFEISKLTEHOEYKIRVCALN 17502  
QY 139 ----- 138  
Db 17503 KVLGEATSVPGTVKPEDKLEAPELDLSELKGIIVRAGGSARIHIPFKGRPMPEITWS 17562  
QY 139 ----- 138  
Db 17563 REEGFTDKVOIEKGVNTQTSIDCNRNDAGKYITLKLENSGGSKSAFVTVKVLDTPGPP 17622  
QY 139 ----- 138  
Db 17623 QNLAVKEVRKDSAFLWEPPIIDGGAKVKNVYIDKRESTRKAYANVSSKCSKTSFKVENL 17682  
QY 139 ----- 138  
Db 17683 TEGAIYYFRVMAENEFVGVPVETVDAYKAAEPSPPGKVTLTDVTSOTSASLMWEKPEHD 17742  
QY 139 ----- 138  
Db 17743 GGSRVLYVVEWQPKGTEKWSIVAESKVCNAVVTGLSSGQEQYQFRVKAYNEKGS DPRVL 17802  
QY 139 ----- 138  
Db 17803 GVPVIAKDLTIQPSLKLFPNTYSIQAGBDLKEIIPVIGRPRNISWVKDGEPLKQTRVN 17862  
QY 139 ----- 138  
Db 17863 VEETATSTVLHIKEGNKDDFGKYTVTATNSAGTATENLSVILEKPPVPGVPRFDEYSA 17922  
QY 139 ----- 138  
Db 17923 DFVVISPEPPAYTGGCQISNVIIVEKRDTTTTWHMVSAIVARTTIKIKLTGTEYQFRI 17982  
QY 139 ----- 138  
Db 17983 FAENRYGKSAPLDSKAVIVQYPFKEPGPTPFVTSISKDQMLVQWHEPVNDGGTKIIGY 18042  
QY 139 ----- 138  
Db 18043 HLEQEKNSILWKLWLNKTPIDQTRFKTTGLDEGEYEFKVAENIVGIGKFSKVECFVA 18102  
QY 139 ----- 138  
Db 18103 RDCDPPGPRPEAIVITRNNVTLKWKKPAYDGGSKITGY IVEKKDLDPGRWMAKSFNTNLE 18162  
QY 139 ----- 138  
Db 18163 TEFTVSGLVEQRYEFVRIARNAAGNFSEPSDSSGAIARDEIDAPNASLDPKYKDVIVV 18222  
QY 139 ----- 138  
Db 18223 HAGETFVLEADIRGKPIPDVWVMSKDGKELEETAARMEIKSTIQKTTLVVKDCIRTDGGQY 18282  
QY 139 ----- 138  
Db 18283 ILKLSNVGGTKSIPITVTVLDRPSPGEPGLKVTGVTAEKCYLANPNPPLQDGGANISHXII 18342  
QY 139 ----- 138  
Db 18343 EKRETSRLSWTQVSTEVQALNXYKVTKLPGNEYPFRVMVAVNKYGIGEPLESGPVTACNPY 18402  
QY 139 ----- 138  
Db 18403 KPPGPPSTPEVSAITKDSMVVTWARPVDDGGTEIEBGYLEKRDKEGVRWTKCNKKTLDL 18462  
QY 139 ----- 148  
Db 18463 RLVRTGLTEGHSEYFRVAENAAAGVGESEPSVFYRACDALYPPGPGSNPKVTDTSRSV 18522  
QY 149 ----- 148

Db 18523 SLAWSKPIYDGGAPVKGVYVEVKEAAADEWTTCTPTGLOGKQFTVTKLKENTENFRIC 18582  
QY 149 -----TLP----- 151  
          |||  
Db 18583 AINSEGVCEPATLPGSVVAQERIEBPPELELDADLRKVVVLRASATLRLFTIKRPEPEV 18642  
QY 152 ----- 151  
Db 18643 KWKAEGILTDRAGIEVTSSTMLVIDNVTRFDSGRYNLTLENNSGSKTAFVNVRLDSP 18702  
QY 152 ----- 151  
Db 18703 SAPVNLTIREVKKDSVTLSEWPEPLIDGAKITNYIVEKRETRKAYATIINCTKTFRI 18762  
QY 152 ----- 151  
Db 18763 ENLQEGCSYFRVLASNEYGIGLPAETTEPVKYSEPPLPGRVTLVDVTRNTATIKWEKP 18822  
QY 152 ----- 151  
Db 18823 ESDGSKITGVVEMOTKGSEKSTCTQVKTLBATISGLTAGEEYVFRVAAVNEKGRSDP 18882  
QY 152 ----- 151  
Db 18883 RQLGVPVIARDIEIKPSVELPFHTFNVKAREQLKIDVPFKGRPOATVNRKDGQTLKET 18942  
QY 152 ----- 151  
Db 18943 RVNVSSKTVTSLSKASKEDVCTYELCVSNSAGSITVPITIVLDRPGPPGPIRIDEV 19002  
QY 152 ----- 151  
Db 19003 SCDISITSMNPPEYDGGCOISNYIVEKKTSTTWHIVSOAVARTSIKIVRLATGSEYQF 19062  
QY 152 ----- 151  
Db 19063 RVCAENRYGKSSYSESAVVAEYFPFPPGPGTPKVVHATKSTMLVTWQVPVNDGGSRI 19122  
QY 152 ----- 151  
Db 19123 GHLEYKERSILWSKANKILIADTVKVSGLDEGLMEYRYVAENIAGIGKSKCEPV 19182  
QY 152 PSRD----- 155  
          |:|  
Db 19183 PARDCDPPGQPEVTNITRSVSLKWSKPHYDGGAKITGIVERRELPGRWLKCNTNI 19242  
QY 156 ----- 155  
Db 19243 QETYFEVTELTEDQRYEFRVFARNAADSVSEPSSESTGPIIIVKDDVEPPRVMMDVKPRDVI 19302  
QY 156 ----- 155  
Db 19303 VRKAGEVLKINADIAGRLPLVISMWKGIEIERARTEIISTDNHTLLTVKDCIRDTGQ 19362  
QY 156 ----- 155  
Db 19363 YVLTLNAGTRSAVNCKVLDKPPGAPGLEINGLTAEKCSLSWGRPOEDGGADIDYXH 19422  
QY 156 ----- 155  
Db 19423 RKKRETSHLAWTICEGELQWTSCKVTLLKGNEYIFRVTGVNKYGVGEPLSVAIKALDP 19482  
QY 156 ----- 155  
Db 19483 FTVPSPPTSLEITSVKESMTLCWSPESDGGSEISGYIIERREKNSLRWRVNRKKPVYD 19542  
QY 156 ----- 155  
Db 19543 LRVKSTGLREGCEYERYVAENAAGLSLPSETSPLIARAEDPVFLPSPSPKIVDSGKTT 19602  
QY 156 ----- 155  
Db 19603 ITIAWVKPLFDGAPITGYTVEYKKSDDDWKTSIQSLRGTEYTTISGLTTGAEYVFRVKS 19662

QY 156 ----- 155  
Db 19663 VNKVGASDPSSDPOIAKEREPEPLFDIDSEMRKTLIVKAGASFTMTVPFRGRPVNVL 19722  
QY 156 ----- 155  
Db 19723 WSKPDTDLRTRAYVDTTDSRTSLTIENANRNDGKYTLTIQNVLSAASLTLLWKVLDTPG 19782  
QY 156 ----- 155  
Db 19783 PPTNITVQDVTKESAVLSWDVPENDGGAPVKNYHIEKREASKAWSVTNNCRLSYKVT 19842  
QY 156 ----- 155  
Db 19843 NLOEGAIYFRVSGENEFVGIPAEWKEGVKITEKPSPEKLGVTISKOSVSLTWLKPE 19902  
QY 156 ----- 155  
Db 19903 HDGSRIVHYVVEALEKGQKNWYKCAVAKSTHHVWSGLRESEYFFRFAENQAGLSDP 19962  
QY 156 ----- 155  
Db 19963 ELLLPVLIKEQLEPPEIDMKNFPSHTVYVRAGSNLKV DIPISGKPLPKVTLSDRGVPLKA 20022  
QY 156 ----- 155  
Db 20023 TMRPNTETAEENLTINLKESVTADAGRYEBITAANSSGTTKAFINIVVLDPRGPPTGPVVI 20082  
QY 156 ----- 155  
Db 20083 SDITEBSVTLKWBPVKYDGGSQVTNYILLKRETSTAVWTEVSATVARTMMKVKMLTGEE 20142  
QY 156 ----- 155  
Db 20143 YQPRIKAENRFGISDHIDSACVTVKLPYTPPGPPSTPWNTVNTRESTITGVHHEPVSNGS 20202  
QY 156 ----- 155  
Db 20203 AVVGYHLEMKDRNSILWQANKLIVIRTHFKVTTISAGLIYEFVRYAENAGVKPSHPS 20262  
QY 156 -----ELTKNOVSL----- 164  
          :::| | |  
Db 20263 EPVLADACEPPRNVRTIDISKNSVLSLWQOPAFDGGSKITGIVERRDLDPGRWTKASF 20322  
QY 165 -----TCL----- 167  
          |||  
Db 20323 TNVTETQFTISGLTQNSQYEFVRFARNAGVSGISNPSEVGPITCIDSYGGPVIDLPLEYT 20382  
QY 168 ----- 167  
Db 20383 EVVKYRAGTSVKLRAGISGKPAPTIEMWKDDKELOTNALVCVENTD LASILIKDADRLN 20442  
QY 168 ----- 167  
Db 20443 SGCYELKLRNAMAASATIRVQILDKPPGPGPIEFKVTAEKITLLWRPPADGGAKIT 20502  
QY 168 ----- 167  
Db 20503 HYIVEKRETSRVVWSMVSEHLEECIITTTKIINGNEYIFRVRVAVNKYGIGEPLESDSVA 20562  
QY 168 ----- 167  
Db 20563 KNAFVTPGPGPIEVTKITKNSMTVVMVSRPIADGGSDISGYFLEKDKKSLGWFKVLKET 20622  
QY 168 ----- 167  
Db 20623 IRDTROKVTGLTENSQYQYRVCVNAAGOGPFSEPFYKAADPIDPPGPPAKIRIADST 20682  
QY 168 ----- 167  
Db 20683 KSSITLGWSKPVYDGSASVTGYVVEITROGEEETWTVSTKGEVTRTEYVYVSNLKPGVNY 20742

QY 168 ----- 167  
Db 20743 FRSAVACAGGEP IEMNEPQAKIDLEAPEIDLDVALRTSVIAKAGEDVQVLIPKGRP 20802  
QY 168 ----- 167  
Db 20803 PPTVTRKDEKNLGS DARYSIENWDSLLTIPQVTRNDTGKXILTIENGVGEPKSTVS 20862  
QY 168 ----- 167  
Db 20863 VKVLDTPAACQKLVKHVSRGTVTLLWDPPLIDGGSPINNVIEKRDATAKRTWSVSHKC 20922  
QY 168 ----- 167  
Db 20923 SSTFKLIDLSEKTPFPFRVLAENEIGIGPCETTEPVKAAVPAPIRDLMSKDSTKTSV 20982  
QY 168 ----- 167  
Db 20983 ILSWTKPDFDGGVITEYVVERKKGQWTHAGISKTCEIYSQLEQSVLEFRFAKN 21042  
QY 168 ----- 167  
Db 21043 EKLSDPVTIGITVRELIIITPEVDSLDIPCAQVTVRIGHNVHLELDYKGPKPSPISWLK 21102  
QY 168 ----- 167  
Db 21103 DGLPKESFVRFSKTENKITLSIKNAKKEHGKYTVILDNAVCRIVAPITVITLGPSPK 21162  
QY 168 ----- 167  
Db 21163 PKGPIRDEIKADSVILSDVPEDNGGGEITCYSIEKRETSQTNWKMVCSSVARTPEKVP 21222  
QY 168 ----- 167  
Db 21223 NLVKDAEQFRVAENRYGVSQPLVSSIIIVAKHQFRIPGPGKPVINYNTSDGMSLTWDA 21282  
QY 168 ----- 172  
Db 21283 PVDGGSEVTGFHVEKKERNLSILQKVNTSPISGREYRATGLVEGLDYQFRVYAENSAGL 21342  
QY 173 --PSD----- 175  
Db 21343 SSPSDPSKFTLAVSPVDPGTPDYIDVRETITLKNPPLRDGSKIVGYSIEKROGNER 21402  
QY 176 ----- 175  
Db 21403 WVRNFTDVSECQYVTGLSPGDRYEFRIIARNNAVGTISPSSSGIIMTRDENVPPIVE 21462  
QY 176 ----- 175  
Db 21463 FGPEYFDGLIIKSGESLRKALVQGRPVPRVTFKDGVEIEKRMNMEITNVLGSTSLFVR 21522  
QY 176 ----- 175  
Db 21523 DATRDHRGVYVEAKNAGSAKAEIKVKQDTPGVVGPFRFTNITGERMTLWMDAPLND 21582  
QY 176 ----- 175  
Db 21583 GCAPITHYIIEKRETSRLAWALIEDKCEAOASYTAIKLINGNEYQFRVSAVNKFGVGRPLD 21642  
QY 176 ----- 175  
Db 21643 SDPVAIQIYTPDAPGIPESNITGNSITLTWARPESDGGSEIQOYILERREKKSTRWV 21702  
QY 176 ----- 175  
Db 21703 KVISKRPISETRFKVTGLTEGNEYEFHVMAENAAAGVPASGISRLIKREFPVNPVPGPPTV 21762  
QY 176 ----- 180  
Db 21763 VKVTDTSKTVSLENSKPVFDGMEIIGYIIECMCKTDLGDWHKVNACVKTXYTVTDLQ 21822  
QY 181 ----- 180

Db 21823 AGEYKFRVSAINGAGKGDSCVETGTIKAVDRLTAPELDIDANFKQTHVVRAGASIRLFI 21882  
QY 181 ----- ESN----- 183  
Db 21883 AYOGRRPTTAVMSKPDNSLSRADIHITDSFSTLTVCNCRNDAGKYTLTVENNSGSKSI 21942  
QY 184 ----- 183  
Db 21943 TFTVKVLDTPCPPGPIITFKDVTGRSATLMDAPILLDGGARLHHYVVEKREASRSWQVIS 22002  
QY 184 ----- 183  
Db 22003 EKCTROIKFVNDLAEGVPYIFRVSANNEYGVGEPEMPEPIVATEQAPPPRLDWDTSK 22062  
QY 184 ----- 183  
Db 22063 SSAVLAWKPDHGGSRITGYLLEMROKGSDLWVEAGHTKQLTFTVERLVEKTEYEFVRK 22122  
QY 184 ----- 183  
Db 22123 AKNDAGYSEPREAPSSVIIKEPQIEPTADLTGITNQLITCKAGSPFTIDVPISGRPAKV 22182  
QY 184 ----- 183  
Db 22183 TWKLEEMRLKETDRVSITTTTKDRTTLTVKDSMRGDSGRYFLLTENTAGTKFTSVTVVVG 22242  
QY 184 ----- GQP----- 186  
Db 22243 RGPVPTGPIEVSSVSAESCVLSWGPEKDGGEITNYIVEKRESGTTAWQLVNSSVRKTQ 22302  
QY 187 ----- 186  
Db 22303 IKVTHLTXYMEYSPRVSSENRFVGVSKPLESAPIIAEHFPVPSPAPTRPEVHVHVSANMSI 22362  
QY 187 ----- ENNYKTT----- 193  
Db 22363 RWBEPYHDGSGKIIGYWEKKERNITLWVKENKVPCLCNCYKVTGLVEGLEQFTVALN 22422  
QY 194 ----- 193  
Db 22423 AAGVSKASEASRPIMAQNPVDAPGRPEVDTRSTVSLINSAPAYDGGKVGYIIEKRP 22482  
QY 194 ----- 193  
Db 22483 VSEVGDGRWLKCNITYIVSDNFFVTALSEGDTYEFVRVLAKNAAGVISKGSESTGPVTCRD 22542  
QY 194 ---pp----- 195  
Db 22543 EYAPPKAELDARLHGLDVTIRAGSDLVLDAAVGGKPEPKIIWTGDKELDICEKVSLOYT 22602  
QY 196 ----- VLDSDG----- 201  
Db 22603 GKRATAVIKFCDRSDSGKYTLTVKNASGTRKAVSYMVVLDSPGCGKLTISRVTQEKCTL 22662  
QY 202 ----- 201  
Db 22663 AWSLPQEDGGAETHYIVERRETSRLNWVIVEGECPTLSYVVTRLIKNNEYIFRVRVANK 22722  
QY 202 ----- SFFL----- 205  
Db 22723 YGPGVPVESEPIVARNSTTIPSPGPIPEEVGTGKEHIIQWTKPESDGGNEISNVLVDKR 22782  
QY 206 ----- 205  
Db 22783 EKESLRTRVKNQYVYDTRLKVTSLMEGCDYQFRVAVNAAGNSEPSESNFISCREPS 22842  
QY 206 ----- 205  
Db 22843 YTPGPPSAPRVDTTKHSISLAWTKPMYDGGTDIVGVYLEMOEKDQWTRVHTNATIRN 22902  
QY 206 ----- 205

Db 22903 TEFTVPLKMGQKYSRVAANVKGMSYSESAIEIEPVERIEIPDELEADLKKVTIR 22962  
QY 206 -----YSLTVDK----- 213  
Db 22963 AGASLRLVSVSGRPPPVITWSKQIDLASRALIDTTESTYSLIIVDKVNRDAGKYTIEA 23022  
QY 214 ----- 213  
Db 23023 -ENQSGKKSATVLVKVYDTGPGPSVKVKEVSRDSVTITWEITPIDGGAPINNVIVEKREA 23082  
QY 214 ----- 213  
Db 23083 AMRAFKVTTKCKSLYRISGLVEGTMHYFRVLPENIYGIGPECETSDAVLSEVPLPA 23142  
QY 214 ----- 213  
Db 23143 KLEVVDVTKSTVTLAWEKPLYDGGSLRTGYVLEACKAGTERWKMVVTLKPTVLEHTVTSI 23202  
QY 214 ----- 213  
Db 23203 NEGEQYLFRIQAQNEKGVSEPRETVTAVTVDQLRVLPTIDLSTMPQKTIHVPAGRPVELV 23262  
QY 214 ----- 213  
Db 23263 IPIAGRPPPAASWFFAGSKLRESERVTVETHKVAKLTIREFTIIRDGTGYTILELKNVTGT 23322  
QY 214 ----- 213  
Db 23323 TSETIKVIILDKPGPTGPIKIDEIDATISITISWPELDDGAPLSGYVVEQDAHRPQW 23382  
QY 214 ----- 213  
Db 23383 LPVSESVRSTRFEKTRLTGNEVYFRVAATNRFGIGYLOQSEVIECRSSIRIPGPETIQ 23442  
QY 214 ----- 213  
Db 23443 IFDVSRDGMTLWYPEDDGGSOVTGYIVERKEVRADRVVRVKNVPTWTRYSRGLTEG 23502  
QY 214 ----- 213  
Db 23503 LEVEHRTAINARGSKPSRPKPIVAMPDIAPPKQPQPRVTDTRTSVSLAWSPEDE 23562  
QY 214 ----- 213  
Db 23563 GGSKVTGYLIEMQKVDQHEWTKCNTPTPKIREYTLTHLPOGAEYFRVLACNAGGPGEPA 23622  
QY 214 -----SRWQO----- 219  
Db 23623 EVPGTVKVTEMLEYDPDYELDERYQEGIFVRQGGVIRLTPIPKGPPFPICKWTKEGODISK 23682  
QY 220 ----- 219  
Db 23683 RAMIATSEHTELVIKEADRGDSGYDVLNKCCKKAYIKVRIGSPNSPEGPLEYDD 23742  
QY 220 ----- 219  
Db 23743 IQVRSVRSWRPPADGGADILGYILERREVPKAAWYITDSRVGTSLVVKGLKENVEYH 23802  
QY 220 ----- 219  
Db 23803 FRVSAENQFISKPLKSEEPVTPKTPLPNPPEPPSNPPEVLDVTKSVSLWSRPKDDGGS 23862  
QY 220 ----- 219  
Db 23863 RVTGYIERKETSTDKVVRHNKTQITTMVTYVTLVDPDAEYQFRIAQNVDVGLSETSPAS 23922  
QY 220 ----- 219  
Db 23923 EPVCKDPDKPSQCELEILSISKDSVTLOWEKEPCDGGKEILGWVVEYRQSGDSAWKK 23982  
QY 220 ----- 219  
Db 23983 SNKERIKQKFTIGLLEATEYFRVFAENETGLSRPRRTAMSIKTLTSGEAPGIRKEM 24042

QY 220 ----- 219  
Db 24043 KDVTTKLGEAAQLSCOIVGRPLPDIKWYRFGKELIOSRKYKMSSDGRTHTLVTWTEQED 24102  
QY 220 ----- 219  
Db 24103 EGVYTCIATNEVEGETSSKLLQLQATPQHPGYPLEKYIYAGVSTLRLHVMYIGRPVPA 24162  
QY 220 -----NVF----- 222  
Db 24163 MTWFHQKLLQNSENITIENTHYTHLVKMNQVORTHAGKYKVLQSNVFGTVDAILDVEI 24222  
QY 223 ----- 222  
Db 24223 QDKPKPTGPVIEALLKNSAVISMKPPADGGSWITNVVEKCEAKEGAEMQLVSSAIS 24282  
QY 223 ----- 222  
Db 24283 VTTCRIVNLTENAGYFRVSAQNTFGISDPLEVSSVYIIKSPFKPGPKPTITAVTKD 24342  
QY 223 SCV----- 226  
Db 24343 SCVANKPPASDGGAKIRNYILEKREKKONKWIISVTTEIRETVFSVKNLIEGLEYEFRV 24402  
QY 227 ----- 226  
Db 24403 KCENLGGESEWSEISEPITTPKSDVPIQAPHKEELNLRVYQSNATLVCKVTGHPKPIV 24462  
QY 227 ----- 226  
Db 24463 KMYROGKEIADGLKYRIOEFKGYHQLIIASVTDATVYQVRATNOGGSVSGTASLEV 24522  
QY 227 ----- 226  
Db 24523 EYPAKIHLPKTLLEGMAVHALRGEVSIKIPFSGRDPVITQKGQDLDNNNGHYQVIYT 24582  
QY 227 ----- 226  
Db 24583 RSFTSILVFPNGVERKDAGFYVVCANRFGIDQKTVELDVADVPDPGRGVKVSASRDSYN 24642  
QY 227 ----- 226  
Db 24643 LTWTEPASDGSKITNYIVEKCATTAEWLRYGOARETRYVINLFGKTSYQFRVIAENK 24702  
QY 227 ----- 226  
Db 24703 FGLSPSEPSEPTITKEDKTRAMNYDEEVDETREVSMTKASHSTKELYEKYMIADLGR 24762  
QY 227 -----MHEA----- 230  
Db 24763 GEGIVHRCVETSSKTYNAKFVKVGTQDVLVKKESILNIAHRNHLHLHESFESMEE 24822  
QY 231 -----LHNH----- 234  
Db 24823 LVNIFEFGIDIFERINTSAFELNEREIVSVYHVQCEALQFLHSHNIGHFDIRPENIY 24882  
QY 235 -----YT----- 236  
Db 24883 QTRRSSTIKIIEFGOARQLKPGDNFRLLFTAPEYYAPEVHQHDVVSVSTATDMNSLGLTVV 24942  
Db 24943 LLSGINPFLAETNQIIEINMNAEYTFDEAFKEISIEAMDFVDRLLVVKERSRMTASEA 25002  
QY 237 -----OXSLSL- 243  
Db 25003 LQHPWLKQKIERVSTKVIRTLKHRRYYHTLIKKDLNMVYSAARISCGGAIRSQKGVSAK 25062  
QY 244 ----- 243  
Db 25063 VKVASTIEIGPVSGQIMHAVGEEGGHVKYVCKIENYDQSTQVTWYFGVROLENSEKEYEITY 25122

```
QY 244 ----- 243
Db 25123 EDGVAILVVKDITKLDGTYRCKVNDYGEDSSYAEFLVKGREVVYCYCRRTMKIKRR 25182
QY 244 ----- 243
Db 25183 TDTMLLERPPEFTPLYNKNTAYGVENRFGVTITVHPPEHVTYKSGQKIKPGDNCKY 25242
QY 244 ----- 243
Db 25243 TPESDKGLYQLINSVTTDDDAEYTVVARNKYGEDSKAKLIVTLHPPTDSTLRPMFKR 25302
QY 244 ----- 243
Db 25303 LLANACQEQSQVCEIRYVGIPTLKWKGQPLSLGNPIETIHEGLDYVALHIDTFL 25362
QY 244 ----- 243
Db 25363 PEDTGYRVTAATAGSTSCQAHQVRLRYKKQEFKSKHEHVRHVQKQIDKTLRMAEIL 25422
QY 244 ----- 243
Db 25423 SGTESVPLTQVAKREALREAAVLYKPAVSTKVKGFRLEIEBKERKLUMPDYDPEPRK 25482
QY 244 -----PGK 246
Db 25483 YKQTTIEEDQRIKQFVPMDSMKWYKIRQYEMPGK 25518
RESULT 15
Q8WZB3
ID Q8WZB3 PRELIMINARY; PRT: 26926 AA.
AC Q8WZB3;
DT 01-MAR-2002 (TremBLrel. 20, Created)
DT 01-MAR-2002 (TremBLrel. 20, Last sequence update)
DT 01-JUN-2002 (TremBLrel. 21, Last annotation update)
DE N2B-titin isoform.
GN TTN.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=20309627; PubMed=10850961;
RA Freiburg A., Trombitas K., Hell W., Cazorla O., Fougereousse F.,
RA Centner T., Kolmerer B., Witt C., Beckmann J.S., Gregorio C.C.,
RA Granzier H., Labelt S.;
RT "Series of exon-skipping events in the elastic spring region of titin
RL as the structural basis for myofibrillar elastic diversity.";
RL Circ. Res. 86:1114-1121(2000).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=21573839; PubMed=11717165;
RA Bang M.L., Centner T., Fornoff F., Geach A.J., Gotthardt M.,
RA McNabb M., Witt C.C., Labelt D., Gregorio C.C., Granzier H.,
RA Labelt S.;
RT "The complete gene sequence of titin, expression of an unusual -700
RT kDa titin isoform and its interaction with obscurin identify a novel
RT Z-line to I-band linking system.";
RL Circ. Res. 89:1065-1072(2001).
DR EMBL; AJ277892; CAD12455.1; .
DR InterPro; IPR000282; Cytochrome receptor_2.
DR InterPro; IPR000719; Euk_pkinase.
DR InterPro; IPR000577; FGyl_kin.
DR InterPro; IPR003961; FN_III.
DR InterPro; IPR001092; HLH_basic.
DR InterPro; IPR003599; Ig.
DR InterPro; IPR003598; Ig_c2.
DR InterPro; IPR003600; Ig_like.
DR InterPro; IPR003006; Ig_MHC.
DR InterPro; IPR002016; Peroxidase.
DR InterPro; IPR002290; Ser_thr_pkinase.
```

```
InterPro; IPR001245; Tyr_pkinase.
Pfam; PF00041; fn3; 132.
Pfam; PF00047; ig; 91.
Pfam; PF00069; pkinase; 1.
ProDom; PD000001; Euk_pkinase; 1.
SMART; SM00060; FN3; 132.
SMART; SM00409; IG; 111.
SMART; SM00408; IGc2; 93.
SMART; SM00410; IG_like; 3.
SMART; SM00220; S_TKc; 1.
SMART; SM00219; TyrKc; 1.
PROSITE; PS00933; FGyl_KINASES_1; UNKNOWN_1.
PROSITE; PS00038; HELIX_LOOP_HELIX; UNKNOWN_1.
PROSITE; PS00290; IG_MHC; UNKNOWN_1.
PROSITE; PS00435; PEROXIDASE_1; UNKNOWN_1.
PROSITE; PS00011; PROTEIN_KINASE_DOM; 1.
PROSITE; PS00109; PROTEIN_KINASE_TYR; UNKNOWN_1.
SQ SEQUENCE 26926 AA; 2993285 MW; 169AB42637A7C1FB CRC64;

Query Match 34.7%; Score 463; DB 4; Length 26926;
Best Local Similarity 0.7%; Pred. No. 0.00021;
Matches 168; Conservative 43; Mismatches 34; Indels 23731; Gaps 53;

QY 2 WLK----- 4
Db 1543 WLKNSDIIVPHKYPKIRIEGKGAALKIDSTVSQDSAMVTATAINKAGRDITTRCKVNVE 1602
QY 5 -----AFYK----- 9
Db 1603 VEFAPEPERKLIIPRGTYRAKEIAAPELEPLHLRYGQEWEGDLYDKEKQKPFKK 1662
QY 10 ----- 9
Db 1663 LLSRLKRFQPAHFECRLTPIGDPTMVVWELHDGCKLEAANRLMINERGCYSLDYGVAY 1722
QY 10 ----- 9
Db 1723 SRDSGIITCRATNKYGTDHTSATLIVKDEKSLVEESQLEPGRKGLQRIEELERMAHEGAL 1782
QY 10 ----- 9
Db 1783 TGVTTDOKEKQKPDIVLYPEPVVLEGETARFCRVTPQPKVNWYLNGLIRKSKRFR 1842
QY 10 -----VAE----- 12
Db 1843 VRYDGHLYLDIVDCKSYDTGEVKVTAENEGVIEHKVLEIQOERDFRSVLRRAPERPE 1902
QY 13 -----KLKEA----- 17
Db 1903 FHVHEPGKLFQEVQKVDPRVDTTETKEVVKLRAERITHEKVPSESELRSKFRKRTGEG 1962
QY 18 ----- 17
Db 1963 YYEAITAVELSKRKDESYEELLRKTDELLHWTKELEKALAEKGKITPTFKPKD 2022
QY 18 ----- 17
Db 2023 IELSPSEAPKIFERIQSQTQVGGSDAHPFRVVRVGGPDCEWYKNGVKIERSDRIYWN 2082
QY 18 ----- 17
Db 2083 PEDNVCVELVRDVTAEADSASIMVKAINIAGETSSHAFLLVQAKLITFTQELQDVVAKER 2142
QY 18 ----- 17
Db 2143 DTMTAFECETSEPFVKVKWKYKGMVEHGDYKRMHSDRKVHFLSILTIPTSDAEDYSCVL 2202
QY 18 ----- 17
Db 2203 VEDENVKTTAKLIVEGAVVEFVKELQDIEVPESYSGELECIIVSPENIEGKWHNDVELKS 2262
QY 18 ----- 17
```

Db 2263 NGKYITRRGRQNLTVKDVTKEDQGEYSFVIDGKKTCTCKLKMKPRPTAILOGLSDOKVC 2322  
Qy 18 ----- 17  
Db 2323 EGDIVQLEVKYLSVESVEGWMKQGEVQPSDRVHIVIDKQSHMLLIEDMTKEDAGNYSFT 2382  
Qy 18 ----- 17  
Db 2383 IPALGLSTGRVSVYSDVITPLKDVNVIETKAVLECKVSPDVTSVKWLNDQIKPD 2442  
Qy 18 ----- 17  
Db 2443 DRVQAIKGTORLIVINRTHASDEGPKLIIVGRVETNLSVEKIKIIRGLRDLTCTETQ 2502  
Qy 18 ----- FMDK----- 21  
Db 2503 NVFVELSHSGIDVLNFKDKEIKPSSKYKIEAHGKIYKLTVLNMMKDDGKYTFYAGE 2562  
Qy 22 ----- 21  
Db 2563 NITSGKLTVAGGAISKPLTDQVAESQBAVFECEVANPDSKGWLRDQKHLPLTNIRSE 2622  
Qy 22 ----- 21  
Db 2623 SDGHKRLIIAATKLDIDIGEITYKVATSKTSAKLKVEAVRIKTKLNLTVTQDAVFTV 2682  
Qy 22 ----- 21  
Db 2683 ELTHPNVKGOWIRKNGVLESNEKYAISVKGTIYSLRIKNCIIVDESIVGPRIGLGASA 2742  
Qy 22 ----- 21  
Db 2743 RLHVETVKI IKKPKDVTALENATVAFESVSHDTPVKWFHKNVEIKPSDKHRLVSEKV 2802  
Qy 22 ----- 21  
Db 2803 HKMLONISPDAGEYTA VVGQLECKAKLFVETLHITKMKNIEVPETKTASFCEVSHF 2862  
Qy 22 ----- 21  
Db 2863 NVPSWLKNGVEIEMSEKFIWQGLHQLIIMNTSTEDSAEYTFVCGNDQVSA TLVTP 2922  
Qy 22 ----- THT----- 24  
Db 2923 IMITSMKLDINAEKDTITFEVTNVEGISYKWLKNGVEIKSTDKCQMRTRKLTSLNIR 2982  
Qy 25 ----- 24  
Db 2983 NVHFGDAADYTFVAGKATSTATLYVEARHIEFRKHIDIKVLEKRAMFECEVSEPDITV 3042  
Qy 25 ----- 24  
Db 3043 QWMKDDQELQITDRIKIQEKYVHRLIIPSTRMSDAGKYTVVAGNVSTAKL FVEGRDVR 3102  
Qy 25 ----- 24  
Db 3103 IRSIKKEVQVIEKQAVVEFEVNEDDVAHWYKDGIEINFQVERHKYVVERRIHRMFIS 3162  
Qy 25 ----- 24  
Db 3163 ETROSDAGEYTFVAGRNRSSVTLVNAPEPPQVLQELQVTVVQSGKPARFCAVISGRPQP 3222  
Qy 25 ----- 24  
Db 3223 KISWYKEQLLSTGFKFLHDGQEYTLILLIEAFPEDAANYTCEAKNDYGVATTSASLSV 3282  
Qy 25 ----- 24  
Db 3283 EVPEVSPDQEMVYPPAIITPLQDVTITSEGPAPFCQVSGTDLKVSWSKDKIKRPSR 3342  
Qy 25 ----- 24  
Db 3343 FFRMTQFEDTVOLEIAEYAPEDEGTYTFVASNAGVQVSSSTANLSLEAPESILHERIEQEI 3402

Qy 25 ----- 24  
Db 3403 EMENKEFSSSFLSAEEGLHSAELQLSKI NETLELLSEPVYSTKFDESEKGTGPIFIKE 3462  
Qy 25 ----- 24  
Db 3463 VSNADISMGDVATLSVTVIGIPKPKIQWFFNGVLLTPSADYKFFVFDGDHSLIILFTKLE 3522  
Qy 25 ----- -cpp----- 27  
Db 3523 DEGEYTCMASNDYKGTICTSAYLKNKSGEGHKDTETESAVAKSLEKLGCGPCPPHFLKELK 3582  
Qy 28 ----- 27  
Db 3583 PIRCAOGLPALFEYTVVGEPAPTVTFWKENQOLCTSVYTTIHNPNGSTGTFVNDPQRED 3642  
Qy 28 ----- CPAPELLG----- -GPSV--- 39  
Db 3643 SGLYICKAENMLGESTCAAELLVLEDDMTDTPCKAKSTPEAPEDFPQPLKGPAAVEAL 3702  
Qy 40 ----- 39  
Db 3703 DSEOEIATFVKDTILKAALITEENOQLSYEHIAKANELSSOLPLGAQLQILSODKLTUP 3762  
Qy 40 ----- FL----- 41  
Db 3763 ESTREFLCINGSIHFPOLKEPSPNLQLOIVOSOKTFSKEGILMPPEPETQAVLSDEKIF 3822  
Qy 42 ----- 41  
Db 3823 PSAMSIQIINSLTVEPLKTLAEPEGNYPOSSIEPPMHSYLTSAEEVLSPEKTVKSDTN 3882  
Qy 42 ----- 41  
Db 3883 REQVRTLQKQBAQASALILSQSLAEHVESLOSPOVMISQVNYEPLVPSEHSTEGGKILI 3942  
Qy 42 ----- FP----- 43  
Db 3943 ESANPLENAGQDSAVRIEEGKSLRFPPLALEEKQVLLKEHSDNVVMPDQIIESKREPA 4002  
Qy 44 ----- 43  
Db 4003 IKKQVEVQGRDLLSKESLLSGIPEEQRLNLKIQICRALQAAVASEQPGFSEWLRNIEKV 4062  
Qy 44 ----- 43  
Db 4063 EVEAVNITQEPRHIMCMYLVTSKASVTEEVTTIIEDVDPQMANLKMELRDALCAIIEEI 4122  
Qy 44 ----- 43  
Db 4123 DILTAEGPRIQOGAKTSLQEEMDSFSGSKVEPITEPEVESKYLISPEVSYFNVQSRVK 4182  
Qy 44 ----- 43  
Db 4183 YLDATPVTKGVASAVSDEKODESLKPSEKESSESSEGTVEEVATVKIQEAEGGFIRBDG 4242  
Qy 44 ----- 43  
Db 4243 PMIHPTLVDTVSEGDIVHLTTSITNAKEVNWYFENKLVPSDEKFKCLQDQNTYTLVIDK 4302  
Qy 44 ----- 43  
Db 4303 VNTEDHQEYVCEALNDSGKTATSAKLTVVKRAAPVIKRIEPLLEVALGHLAKFTCEIQS 4362  
Qy 44 ----- 43  
Db 4363 APNVRFQWFKAGREIYESDKSIRSKYISSLEILRTQVDCGEYTCCKASNEYSVSCSTA 4422  
Qy 44 ----- PKPKDTLMI----- 52  
Db 4423 TLTVTVPGEKKVKRLLLPKRPKEPEVVLKSVLRKRPEEPEKPKLEKKVPAVPE 4482

QY 53 ----- 52  
Db 4483 PPPPKPVEEVTVTKRERKIPETKVPKPAIPAPKPKPEAEVKTIKPPPPVEP 4542  
QY 53 -----SRTP----- 56  
Db 4543 EPTPIAAPVTVPVVGKAKAKAPKEAAKPKGPIKGVPKKTSPIEARRKLRPGSGEK 4602  
QY 57 ----- 56  
Db 4603 PPDEAPFTYQKAVPLKFKVKEIKDIIITSEFVGSAAIFECLVSPSTAITTWKMGDSNIR 4662  
QY 57 -----EVTVC----- 61  
Db 4663 ESPKHFRIADGDKRLHIIDVQLSDAGEYTCVLRGNKEKTSFPAKLVEELPVRFVKYLE 4722  
QY 62 ----- 61  
Db 4723 EEVTVVGQPLYLSCELNKERDYYVRKDGKIYVEKPGRIYVGVIGLMRALTIINDADDTDA 4782  
QY 62 ----- 61  
Db 4783 GTYTVTVENANNLECSVCVVVEVIRDLVKPIRDQHVKPGTAIPACDIADKTPNIKWF 4842  
QY 62 ----- 61  
Db 4843 KGYDEIPAEPNDKTEILRDGNHLYLKIKNAMPEDIAEYAVEIBGKRYPAKLTLGEREVEL 4902  
QY 62 ----- 61  
Db 4903 LKPIEDVTIYKESAFDAISEADIPGQWKLKCELLRSPPTCEIKAEGGKRFLLRKVK 4962  
QY 62 ----- 61  
Db 4963 LDQAGEVLYQALNAITAILTVKEIELDFAVPLKDVTVPERQARFECVLTREANVIWSK 5022  
QY 62 ----- 61  
Db 5023 GPDIIKSSDKFDIIADGKKHILVINDSQFDEGVYTAEVGKTSARLFTVGTIRLKFMS 5082  
QY 62 -----VVDVSHE----- 68  
Db 5083 LEDQTVKEGATAFVCELSHEKHMVVMFKNDAKLHTSRTVLISSEKTHKLEMKEVTLDD 5142  
QY 69 ----- 68  
Db 5143 ISQIKAQVKELSTAQIKVLEADPYFTVKLHDKTAVEKDEIILKCEVSKDVPVKWFKDGE 5202  
QY 69 ----- 68  
Db 5203 EIVSPKYSIKADGLRILKIKKADLKDKGEYVDCGTDKTKANVTVEARLIKVEKPLYG 5262  
QY 69 ----- 68  
Db 5263 VEVVGETAHEIELSEPDVHGQWKLKGOPLTASPDCEIIEDGKHHILHNCQLGWTGE 5322  
QY 69 ----- 68  
Db 5323 VSPQANAASANKVKELPLIFITPLSDVKVFEKDEAKFECEVSREP KTFRWLKGTOEI 5382  
QY 69 ----- 68  
Db 5383 TGDREFELIKGTGKHSWIKSAAFEDEAKYMEADKHTSGKLIIEGIRLKFUTPLKDVT 5442  
QY 69 ----- 68  
Db 5443 AKEKSAVFTVELSHDNIYVKFNQDLHTTRSVSMQDEGKTHSITFKDLSIDDTSQIR 5502  
QY 69 -----DP----- 70  
Db 5503 VEAMGMSSEAKLTVLEGDPTFTGKLODYGVEKDEVELOCEISKADAPVKWFKDGKEIRP 5562  
QY 71 ----- 70

Db 5563 SKNAVIKADGKKRMLILKKALKSDIGQYTCDCGTDKTSGLKDIEDREIKLVRPLHSVEVM 5622  
QY 71 ----- 70  
Db 5623 ETETARFETEISEDDIHANWKLKGEALLQTPDCEIKEEGKIHSILVHLNCRDQTGGVDFQ 5682  
QY 71 ----- 70  
Db 5683 AANVSSAHLRVKPRVIGILLRPLKDVTVTAGETATFDCELSYEDIPVEWYLKGGKLEPSD 5742  
QY 71 ----- 70  
Db 5743 KVVPRSEGVHTLTLRDVKLEDAGEYQLTAKDFKTHANLFVKBPPEFTKPLEDQTVVEEG 5802  
QY 71 ----- 70  
Db 5803 ATAVLECEVSRENKAKWFKNGTEILKSKYEIVADGRVKLVIHDCPTEDIKTYTODAK 5862  
QY 71 ----- 70  
Db 5863 DFKTSCNLNVVPPHVEFLRPLTDLQVREKEMARFCELSRENKAKWFKDGAELKGGKY 5922  
QY 71 ----- 70  
Db 5923 DIISGAVRILVINKLLDDEAEYSCEVTRTARTSGMLTVLEEEAVFTKNLANIEVSETDT 5982  
QY 71 ----- 70  
Db 5983 IKLVEVSKPGAENVYWKGDDEIETGRYEILTEGRRILVIONAHLEDAGNINCRLPSS 6042  
QY 71 ----- 70  
Db 6043 RTDGKVKVHELAAEFISKPNLEILEGEKAEFVCSISKESFPVQWKRDDTLSEGDYDV 6102  
QY 71 -----EVKF 74  
Db 6103 IADGKKRVLVKDATLQDMGTYVMVGAARAAHLTVIEKLRIVVPLKTRVKEQEVWF 6162  
QY 75 N----- 75  
Db 6163 NCEVNTGAKAKWFRNEEAIFDSSKYIILQKDLVYTLRIRDAHLDDQANYNVSLTNHRGE 6222  
QY 76 ----- 75  
Db 6223 NVKSAANLIVEEDLRIVEPLKDIETMEKKSVTFWCKVNRNLNVLTKWTKNGEEVFPDNRV 6282  
QY 76 ----- 75  
Db 6283 SYRVDKYKHLTIKDCGFPDEGEYIVTAGQDKSVAELLIIEAPTEFVEHLEDQTVTEFDD 6342  
QY 76 ----- 75  
Db 6343 AVFSCOLSREKANVKWYRNGREITEGKKYKFEKDGSTHRLIIKDCRLDDECEYACGVEDR 6402  
QY 76 ----- 75  
Db 6403 KSRARLFVEEIPVEIIRPPODILEAPGADVVFVLAELNKKDKVEVQWLRNNVWVVGDKHQ 6462  
QY 76 ----- 75  
Db 6463 MSEGKTHRLQICDIKPRDQGEYRFAKDKEARAKELAAAPKIKTADQLDVLVDVGKPLTM 6522  
QY 76 ----- 75  
Db 6523 VVPYDAYKRAEAEWFKENEPLSTKTIDTTAEQTSFRILEAKGDKGRYKIVLQNKHGKAE 6582  
QY 76 ----- 75  
Db 6583 GFINKLVIDVPGVPRNLEVTETFDGEVSLAWEPEPLTDGSGKIIGYVYVRRERRDKRTWVLA 6642  
QY 76 ----- 75

Db	6643	TDRAESCEFTVTLQKGVVEYLFVRVSARNRVGTGEPVETDNPVEARSKYDVGGPPLNWTI	6700
Qy	76	-----	75
Db	6703	TDVNRFGVSLTWEPPEYDGGAEITNVYIELRDKTSIRWDTAMTVRAEDLSATVDVVEGQ	6762
Qy	76	-----	75
Db	6763	EYSFRVRAQNRIGVGKPSAATPFVKVADPIERPSPVNLTSDDQTQSSVOLKWEPLPKDG	6822
Qy	76	-----	75
Db	6823	GSPILGYIIBERCEGKDNWIRCNMKLVPELTYKVTGLEKGNKYLVRSAENKAGVSDPSE	6882
Qy	76	-----	75
Db	6883	ILGPLTADDAFVEPTMDLSAFKDGLEVIPNPITILVPSTGYPRPTATWCFGDKVLETGD	6942
Qy	76	-----	75
Db	6943	RVKMKTL SAYAELVISPSERSDKGYTLKLENRVKTI SGEIDVNVVIARPSAPKELKFEDI	7002
Qy	76	-----	75
Db	7003	TKDSVHLTWEPDDDDGGSPLTGYVVEKREVSRTWTVMDFVTDLEFTVPLVQKLEYLF	7062
Qy	76	-----	75
Db	7063	KVCARNKCGPEYAYDEPNVMSTPATVPDPPEYKWRDRTANSIFLTWPPKNDGSGRI	7122
Qy	76	-----	75
Db	7123	KGYIVERCPRGSKDWACGPEVAETKMEVTGLECKGYAYRVKALNRQGASKPSRPTEEI	7182
Qy	76	-----	75
Db	7183	QAVDQEAPEIFLDVKLLAGLTAKGTIELPATVTGKPEPKITWTKADMILKODKRITI	7242
Qy	76	-----	75
Db	7243	ENVPKKTVTIVDSKRSDDTGTVII EAVNVCGRATAVEVNVLDKPGPPAADFIDVTNES	7302
Qy	76	-----	75
Db	7303	CLLTWNPPRDDGGSKITNVWERRATDSEVWHKLSSTVKOTNFKATKLIPNKYEIRVAA	7362
Qy	76	-----	75
Db	7363	ENMYGVEPVOASPITAKYQFPDPGPPTRLEPSDITKDAVTLTWCEPDDDDGGSPITGYVW	7422
Qy	76	-----	75
Db	7423	ERLDPDTDKWVRCKMPVKDTTYRVKGLTNKKYRFRVLAENLAGPGKPSKSTEPILIKD	7482
Qy	76	-----	75
Db	7483	PIDPPMPGKPTVKOVGKTSVRLNWTKEHGGAKIESYVIEMLKTGTDEWVRVAEGVPT	7542
Qy	76	-----	75
Db	7543	TQHLLPCLMEGOEYSFRVRVAVNKAGESEBSPDVLCKREKLYPPSPRWLEVINITKNT	7602
Qy	76	-----	75
Db	7603	ADLKWTVPEKDGGSPITNYYIVEKRDVRRKGQWTDVTYKDTKCTVTPLTEGSLYVFRVAA	7662
Qy	76	-----	75
Db	7663	ENAIQSDYTEIEDSVLAKDFTTTPGPPYALAVDVTKRHHVLDLKWEPKNDGGRPIQRYV	7722
Qy	76	-----	75
Db	7723	IEKKERLGRTRWKAGTKTAGPCDNFRVTDVIEGTVEQFOVRAENEAAGVGHSEPTILSIE	7782

Qy	76	-----75	-----75
Db	7783	DTPSPPLDLHVTDAGRKHIAIAWKPPEKNGSGPIIGYHVEMCPVGTETKWMVRNSRPI	7842
Qy	76	-----75	-----75
Db	7843	KDLKFVBEGVDPKKEYVLRVRAVNAIGVSEIPSEISENVVAKDPDCKPTIDLETHDIIVI	7902
Qy	76	-----WYVDG-----80	-----80
Db	7903	EGEKLSIPVFRAPVPVTVSHKDGKEVKASDRLTMKNDHISAHLEVPKSVRADAGIYTI	7962
Qy	81	-----80	-----80
Db	7963	TLENKLSATASINVKVIGLPGCKDIKASDITKSSCKLTWPEPEFDGCTPILHYVLERR	8022
Qy	81	-----80	-----80
Db	8023	EACRRYIIPVMSGENKLSWTVDLIPNGEYFFRVKAVNKVGGGEYIELKNPVIADQPQP	8082
Qy	81	-----VEVHNA-----86	-----86
Db	8083	PDPVDEVHNPTAEAMTITWKRPPLYDGGSKIMGYIIEIKAEERWKRCNHLVPILTY	8142
Qy	87	-----86	-----86
Db	8143	TAKGLEEGKEYQFRVRAENNAAGISEPSRATPTTKAVDPTDAPKVIILRTSLEVKRGDEIAL	8202
Qy	87	-----86	-----86
Db	8203	DASISGSPYTIWIKDENVIPEEIKKRAAPLVRRRKGEVQEEFPVLPTQRLSIDNS	8262
Qy	87	-----86	-----86
Db	8263	KKGESQLRVDSLDPDHGLYMIKVENHDGIAKAPCTVSLDTPGPPINFVEDIRKTSVL	8322
Qy	87	-----KTKP-----90	-----90
Db	8323	CKWEPLDDGGSEIINYTLEKKDKTKPOSEWIVVTSTLRHCKYSVTKLIEGKEYLFRVRA	8382
Qy	91	-----90	-----90
Db	8383	ENRFGPPPCVSKPLVAKDPFGPDAPDKPIVEDVTSNMLVKWNEPKONGSPILGYWLE	8442
Qy	91	-----90	-----90
Db	8443	KREVNSTHWSRVNKSLLNALKANVDGLLEGLTYVFRVCAENAAAGPKFSPDPKTAHDP	8502
Qy	91	-----90	-----90
Db	8503	ISPPGPPIPRVTDTSSTTIELEPPAFNGGGEIVGYFVDKQLVGTNEWSRCETEKIKVR	8562
Qy	91	-----90	-----90
Db	8563	QYTVKEIREGADYKLRVSAVNAAGEPGGETQPTVVAEPQEPAPVELDVSVKGGIOMAG	8622
Qy	91	-----90	-----90
Db	8623	KTLRIPAVVTGRPVTKVWTKKEGELDKDRVVDNVGTSKSELIINDALRKDHCRYITAT	8682
Qy	91	-----90	-----90
Db	8683	NSGSKFAARVEVFDVPGVLDLKPVTNTRKWCLLNWSDDPDGSGEITGFIERKDAK	8742
Qy	91	-----90	-----90
Db	8743	MHTWRQPIETERSKCDITCLLEBQEYKFRVIAKNKFGCGPPVEIGPILAVDPLGPPTSPE	8802
Qy	91	-----90	-----90
Db	8803	RLTYTERTKSTITLDWKPEPRNSGGSPIOGYIIEKRRHDKDPFERVNRKLCPTTSFLVENL	8862



QY 91 ----- 90  
Db 8863 DEHQYEFVRKAVNEIGESPSPLNVVIOQDDEVPTIKLRLSVRGDTIKVKAGEPVHIP 8922  
QY 91 ----- 90  
Db 8923 ADVTGLPMPKIEMSKNETVIEKPTDALQITKEEVSSEAKTELSIPKAVREDKGTYTVA 8982  
QY 91 ----- 90  
Db 8983 SNRLGSVFRNVHVEYDRPSPRNLAVTDIKAESCYLTDAPLDNGGSEITHYVIDKRDA 9042  
QY 91 ----- 90  
Db 9043 SRKKAEEVNTAVEXRYGIWKLIPNGQYEFVRVRAVNKYGISDECKSDKVVIQDPYRLP 9102  
QY 91 ----- 90  
Db 9103 GPPGPKVLARTKGSMLVSWTPPLDNGGSPITGYWLEKREEGSPYWSRAPITKYGLK 9162  
QY 91 ----- 90  
Db 9163 GVEENVPRLLGVKYQFRAMAINAAGIGPPSESDPEVAGDPIFPFGPPSPCEVKDKTKS 9222  
QY 91 ----- 90  
Db 9223 SISLGWKPAPKAGGSPIKGYIVEMQEEGTTDWKRVNEPDKLITTCCECVWPNLKELRKYR 9282  
QY 91 ----- 90  
Db 9283 RVKAVNEAGESEPSDTTGEIPATDIOBEVEFIDIGAQCLVCKAGSQIRIPAVIKGRPT 9342  
QY 91 ----- 90  
Db 9343 PKSSWEFDGKAKKAMKDCVHDIPEDAQLETAENSSVIIPECKRSHTGKYSITAKNKAGQ 9402  
QY 91 ----- 90  
Db 9403 KTANCRVKVMDVPGPKDLKVSIDITRGSRLSWKMPDDGGDRIGYVIEKRTIDGRAWT 9462  
QY 91 -----REQY----- 95  
Db 9463 KVNPDGSTTFVVDLLSEQQYFFRVRARENRFGIGPVETIQTARTDPIYPPDPPIKLK 9522  
QY 96 ----- 95  
Db 9523 IGLITKNTVHLSWPKPKNDGGSPVTHIVECLAWDPTGTKEAWRQCNKRDRVEELQFTVE 9582  
QY 96 ----- 95  
Db 9583 DLVEGGEYFRKAVNAAGVSKPSATVGPCDCORPDPSPSIDLKEFMEVEEGTNVIVAK 9642  
QY 96 ----- 95  
Db 9643 IKGVPFPTLWFKAPPKPONKEPVLVDTHVNLVVDVDTCTLVIPQSRSSDTGLYTITAV 9702  
QY 96 ----- 95  
Db 9703 NNLTASKEMRLNVLRGPPVGPPIKFEVSADQMTLSWPPPKDDGGSKITNVIKREA 9762  
QY 96 ----- 95  
Db 9763 NRKTWVHVSSEPECTYTIPLKLEGHEVYFRIMAQNKYIGICEPLDSEPTARNLFSVPGA 9822  
QY 96 ----- 95  
Db 9823 PDKPTVSSVTRNSMTVNWEEPEYDGGSPVTGYWLEMKDOTTSKRWKVRNRDPKAMTLGVS 9882  
QY 96 ----- 95  
Db 9883 YKVTGLIEGSDYQFRVAINAAGVPASLPSPDAPARDPIAPGPPFPKVTDTWKSSADL 9942  
QY 96 ----- 95

Db 9943 EWSPLKDGGSKVTVGYIVEYKEGKEBWKGDKEVRGTLVVTGLKEGAFYKFRVSAVN 10002  
QY 96 ----- 95  
Db 10003 IAGIGEPGEVTDVJEMKDRLVSPDLQLDASVRDRIVVHAGGVIRIIAYVSGKPPPTVTWN 10062  
QY 96 ----- 95  
Db 10063 MNERTLPQEAETIETTAISSVMVKNCORSHOGVYSLAKNAGEBKKTIIIVDVLDPGPV 10122  
QY 96 ----- 95  
Db 10123 GTPFLAHLNLTNESCKLTWFSPEDDGGSPITNVIKRESRRAMTPTVYTTVTRONATVQG 10182  
QY 96 ----- 95  
Db 10183 LIOGKAYFFRIAENSIGMGFFVETSEALVIREPITVPERPEDLEVKEVTNTVTLTWNP 10242  
QY 96 ----- 95  
Db 10243 PKYDGSSEIINYVLESRLIGTEKFHKVTNDNLLSRKYTVKGLKEGDTYEYRVSANVIGQ 10302  
QY 96 ----- 95  
Db 10303 GKPSFCTKPTCKDELAPPTLHLDFRDKLTIRVGEAFALTGRYSGKPKPKVSWFKDEADV 10362  
QY 96 -----NST----- 98  
Db 10363 LEDDRTHIKTTPATLALAKIKAKRSDSGKYCVVVENSTGSRKGCQNVNVDPPGPPGPV 10422  
QY 99 ----- 98  
Db 10423 SFDEVTKDYVVISWKPPDLDDGGSKITNYIIIEKVEGKDVWMPVTSASAKTCKVSKLLEG 10482  
QY 99 ----- 98  
Db 10483 KDYIFRIHAENLYGISDPLVSDSMKAKDRFRVPDAPDQPIVTEVTKDSALVTWNRKPHDGG 10542  
QY 99 ----- 98  
Db 10543 KPITNYILEKRETMKRWARTKDPHYPYTKFRVPDLEGGQYEFVRSAENEIGIDPSP 10602  
QY 99 ----- 98  
Db 10603 PSKPVFAKDPTAKSPPVNPEAIDTTCNSVDLTWQPPRHDGGSKILGYIVEYQKVGDEW 10662  
QY 99 -----YRVVSV----- 104  
Db 10663 RRANHTPESCPETKYKVTGLRDGQTKYFRVLAVNAAGESDPAHVPEPVLVKDRLEPPELI 10722  
QY 105 -----LTV----- 104  
Db 10723 LDANMAREHIKVGDTLRLSAIKGVPPPKVTWKEDRDAPTKARIDVTPVGSKLEIRNA 10782  
QY 105 -----LTV----- 107  
Db 10783 AHEDGGIYSLIVENPAGSKTVSVKVLVDKPGPPRDLLEVSEIRKDCSYLTWKPELDGGS 10842  
QY 108 -----LH----- 107  
Db 10843 VITNVVVERRDVAQAQWSPLSATSKKKSHFAKHLNEGNYQLFRVAENQYGRGPFVETPK 10902  
QY 108 -----LH-----ODWL-- 113  
Db 10903 PIKALDPLHPPGPKDLHHVDVKTEVSLVWNNKPDORGSPITGYLVEYQEGTDWIKF 10962  
QY 114 ----- 113  
Db 10963 KVTNLECVTGLQOGKTYRFRVRAENIVGLPDTTPIECQEKLVPPSVELDVKLEIG 11022  
QY 114 ----- 113

Db 11023 LVVAGTIVRFPAIRGVVPVPTAKWTTDGGSEIKTDEHYTETDNFSSVLTIKNCLRRDTG 11082  
QY 114 ----- 113  
Db 11083 EYQITVSNAGSKTAVHLTVLDVPGPPGTINILDTPEHMTISWQPKDDGGSPVINY 11142  
QY 114 ----- 113  
Db 11143 IVEKQDTRKDTWGVSSGSKTKLIPHLOKGEYVFRVAENKIGVPPLDSTPTVAKH 11202  
QY 114 ----- 113  
Db 11203 KFSPPSPGKVVVTDITENAAVSWTLPKSDGSGPITGYMERREVTGKWRVKNKTPID 11262  
QY 114 ----- 113  
Db 11263 LKFRVTGLYEGNTYEFVRFAENLAGLSPSSDPIKACRPKPGPPINPKDKSRET 11322  
QY 114 ----- 113  
Db 11323 ADLWTKPLSDGSGPILGYVEQCPGTAQWNRINKDELIROCAFVPGCLIEGNEYRPRI 11382  
QY 114 ----- 113  
Db 11383 KAANIVGEGEPRELAESVIAKDILHPPEVELDVTCRDVITVRVGTIRILARVKGREPED 11442  
QY 114 ----- 113  
Db 11443 ITWKEGVLVREKRVLDLIQDLPRVELOIKEAVRADHGKIISAKNSSGHAOGSAIVNVL 11502  
QY 114 ----- 113  
Db 11503 DRPGPQNLKVTNVTKENCTISWENPLDNGSEITNFIVEYRKPNKGWSIVASDVTKRL 11562  
QY 114 ----- 113  
Db 11563 IKANLLANNEYFRVCAENKVGVGPTIETKTPILAINPIDRPGEPENLHIADKGTFFVL 11622  
QY 114 ----- 113  
Db 11623 KWRRPDYDGGSPNLSVHVERRLKGDDWERVHKSGIKETHYMWDRVCVENQIYEFVRVQTKN 11682  
QY 114 ----- 113  
Db 11683 EGESDWWKTEEVVVKEDLQKPVLDLKLGSVLTVRAGDTIRLEAGVRGKPFPEVAWTKDK 11742  
QY 114 ----- 113  
Db 11743 DATDLTRSPRVKIDTRADSSKESLTAKKRSDDGKYVVTATNTAGSFVAYATVNVLDKPGP 11802  
QY 114 ----- 113  
Db 11803 VRNLKIVDVSSDRCTVCWDPEDDGGCEIQNYILEKCEKTRVWSTYSATVLTPGTIVTR 11862  
QY 114 ----- 113  
Db 11863 LIEGNEYIFRVAENKIGTPTESKPVIAKTKYDKPGRPDPPEVTKVSKEMTVVWNP 11922  
QY 114 ----- 113  
Db 11923 EYDGGKSITGYFLEKKEKHSRWRVPVKNKSAIPERRMKVQNLPLPDHEYQFRVKAENEIGIG 11982  
QY 114 ----- 113  
Db 11983 EPSLPSRPVAKDPIEPPGPTNFRVDTTKHSITLWGKPKVPYDGGAPILIGYVWEMRPKI 12042  
QY 114 ----- 113  
Db 12043 ADASPDGKRCNAAAQVLRKEFTVTSLDENOEYFRVCAQNOVGIGRPAELKEAIKPE 12102  
QY 114 ----- 113  
Db 12103 ILEPPEIDLASMRKLVIVRACCPILFAIVRGPAPKVTKVRKVGIDNVVRKGGQVLDVDT 12162

QY 114 ----- 113  
Db 12163 MAFLVIPNSTRDDSGKYSLTLVNPAGEKAVFVNVVRVLDTPGPPVSDLVKSDVTKTTSCHVSW 12222  
QY 114 ----- 113  
Db 12223 APPENDGGSQVTHYIVEKREADRKTWSTVTPEVKKTSFHVNTLVPCNEYIFRVTAVNEVG 12282  
QY 114 ----- 113  
Db 12283 PGVPTDVPKPVILASDPLSEDPDRPKLEVTEMTKNSATLAWPLRDLRGGAKIDGYITSYRE 12342  
QY 114 ----- 113  
Db 12343 ESOPADRWTEYSVVVXKDLVLVTGLKEGKIKFVRVAARNAVGVSLPREAGVYEAKEQLLP 12402  
QY 114 ----- 113  
Db 12403 PKILMEQITIKAGKKLRIEAAHVYKGPHTCKKKGEDEVVTSSEHLAVHKADSSSLITIK 12462  
QY 114 ----- 113  
Db 12463 DVTRKDSGYSLTAENSSGTDTKIKVVVMDAPGPPQPPFDISDADACSLSWHIPLEBD 12522  
QY 114 ----- 113  
Db 12523 GGSNITNIVEKCDVSRGDMVTALASVTIKTSRVGKLIPGOEYIFRVAENRFGISEPLT 12582  
QY 114 ----- 113  
Db 12583 SPKMQVAFPGFVPSPEPKNARVTKNKDCIFVANDRPDSGGSPIIIGYLIERKERNSLWV 12642  
QY 114 ----- 113  
Db 12643 KANDTLVRSTEYPCAGLVEGLEYSFRIYALNAGSSPKPTEYVYVTAARMVDPPOKPEVI 12702  
QY 114 ----- 113  
Db 12703 DVTKSTVSLIWARPKHDGSKIIGYFVEACKLPDGKWRCNTAPHQIQOEYVATATGLEBK 12762  
QY 114 ----- 113  
Db 12763 AOYQFRAIARTAVNISPPSEPSDPTVILAENVPRIIDLVSAMKSLTIVKAGTNVCLDATV 12822  
QY 114 ----- 113  
Db 12823 FGKPMPTVSWKKGDTLLKPAEGIKMAMORNLCLELFSVNRKDCDGYTTITAENSSGSKSA 12882  
QY 114 ----- 113  
Db 12883 TIKLAVLDKPGPPASVKINKMYSDRAMLSWEPPELGGSEITNYIVDKRETSRPNNAQVS 12942  
QY 114 ----- 113  
Db 12943 ATVPIITSGVEKLIBGHEYQFRICAENKYGVGDPVFTEPAIAKNPYDPPGRCDPPVISNI 13002  
QY 114 ----- 113  
Db 13003 TKDHMTVSWKPPADGGSPITGYLLEKRETOAVNWKVNRKPIERTLKATGLQEGTEYE 13062  
QY 114 ----- 113  
Db 13063 FRVTAINKAGPKPSKASAAVARDPOYPAPPAPPKVYDTRSSVSLSWGKPAYDGGSP 13122  
QY 114 ----- 113  
Db 13123 IIGYLVEYKRADSDNWRNCLPNQLQKTRFEVTGLMEDTOYQFRVYAVNKGIGYSDPSDVP 13182  
QY 114 ----- 113  
Db 13183 DKHYPKDILIPPEGELDADLRKTLILRAGVTMRILYVPVKGRRPPKITWSKPNVNLDRDIG 13242

QY 114 ----- 113  
Db 13243 LDIKSTDFTLRCENVNKYDAGKYLITLNSCGKEYTIIVKVLDTGPPVNVVTKRIS 13302  
QY 114 ----- 113  
Db 13303 KDSAYVTWEPIIDGGSPIINVVQKRDAERKSWSTVTECSKTSFRVANLECKSYFFR 13362  
QY 114 ----- 113  
Db 13363 VFAENEYIGDPGETRDAVKASQTPGPVVDLKVRVSXKSSCSIGWKPHSDGGSRIIGYV 13422  
QY 114 -----NGKEY----- 118  
| | | |  
Db 13423 VDFLTEENKQORVMKSLSLQYSAKDLTEGKEYTFRVSAENENGCTPSEITVVARDOWA 13482  
QY 119 ----- 118  
Db 13483 PDLDLGLDLYLAKENSFRLKIPKGPAPSVSWKKGEDPLATDTRVSESSAVNTT 13542  
QY 119 ----- 118  
Db 13543 LIVYDCKSDAGKYTITLKNVAGTKEGTISIKVVGKPGIPTGPIKPFDEVTAEAMTLKWP 13602  
QY 119 ----- 118  
Db 13603 PKDDGSEITNYILEKRDSVNNKWTCASAVQKTTFRVTRLHEGMEYTFRVAENKYGVG 13662  
QY 119 ----- 118  
Db 13663 EGLSEPIVARHPFDPDAPPPNIVDVHRDSVLTWTDPKKGGSPITGHHLEKBERNS 13722  
QY 119 ----- 118  
Db 13723 LLWRANKTPIMRDFKVTGLTEGLEYEFRVMAINLAGVGKPSLPSEPVVALDIDPPGK 13782  
QY 119 ----- 118  
Db 13783 PEVINITRNSVTLIWTEPKYDGGHKLTYIVEKRDLPSKSMKANHVNPBCAFTVTDLV 13842  
QY 119 ----- 118  
Db 13843 EGGKYEFRIIRAKTAGAISAPSESTETIICKDEYEAPTIVLDPTIKDGLTIKAGDTIVLN 13902  
QY 119 ----- 118  
Db 13903 AISILGKPLPKSSWSKAGDIRPSDITQITSTPTSSMLTIKYATRKDAGEYTIATNPF 13962  
QY 119 ----- 118  
Db 13963 TKVEHVKTVDVPPGPPVEISNVAEKATLTWTPLEDGGSPIKSVILEKRETSRLW 14022  
QY 119 ----- 118  
Db 14023 TVVSEDIQSCRHVATKLIQGNEXIFRVSANHYKGEPVQSEPVKMVDRFGPPGPEKPE 14082  
QY 119 -----KCKVSN--- 124  
: | | :  
Db 14083 VSNVTKNATVSWKRPVDDGCGSEITGYHVERREKSLRWVRAIKTPVSDLRCKVTGLQEG 14142  
QY 125 ----- 124  
Db 14143 STYEFVSAENRAGIPGPSEASDVLMDAAYPGPSPNPHVTDTKKSASLANGKPHYD 14202  
QY 125 ----- 124  
Db 14203 GGLEITGVVEHQVGEAWIKDTGTALRITQFVVPDLQTKENFRISAINDAGVGE 14262  
QY 125 -----KALPAP----- 130  
| | | |  
Db 14263 AVIPDVEIVEREMAPDELDLRLTLVVRAGLSIRIFVPIKGRPAPEVWTKNINLKN 14322  
QY 131 ----- 130

Db 14323 RANIENTESEFTLLIPECNRYDTGKFVMTIENPAGKKSGFVNVRLDTPGPVLNLRPTDI 14382  
QY 131 -----IEK----- 133  
: | |  
Db 14383 TKDSVTLHNDLPLIDGGSRTNYIVEKREATRKSYSATATTCHKCTYKVTGLSEGCYFF 14442  
QY 134 ----- 133  
Db 14443 RVMAENEYIGEPTETTEPVKASEAPSPDLSINIMDITKSTVSLAWPKPKHGGSKITGY 14502  
QY 134 ----- 133  
Db 14503 VIEAQRKGDOWHTITTVKGLECVVRNLTEGEYTFQYMAVNSAGRSAPRESRPVIVREQ 14562  
QY 134 ----- 133  
Db 14563 TMLPELDLRGIYQKLVIAKAGDNIKVEIPVLGRPKPTVWKKGDOIILKQTORVNFETTAT 14622  
QY 134 ----- 133  
Db 14623 STILNINECVRSDSGPYPLTARNIVGEVDVITIQVHDI PGPTGPIKFEVSSDEVTF 14682  
QY 134 ----- 133  
Db 14683 WDPPENDGGVPISNVVEMROTDSTTWVELATTVIRTTYKATRLTTGLEYOFRVKAQNR 14742  
QY 134 ----- 133  
Db 14743 GVGPGITSACIVANYPFKVPGPPTQVTAVTKDSMTISWHEPLSDGSPILGYHVERKE 14802  
QY 134 -----TISKA----- 138  
| : | | |  
Db 14803 RNLGWLQVSKALVPGNIFKSSGLTDGIAYEFVIAENMAGSKPSKPSKSEPMALDPTDP 14862  
QY 139 ----- 138  
Db 14863 PGKVPNLNTRHTVTLKWKPEYTGKFIITSYIVEKRDLPNGRWLKNFNSILENEFTVS 14922  
QY 139 ----- 138  
Db 14923 GLTEDAAEFVRIAANAAGISPPSEPSDAITCRDDVEAPKIKVDVKFKDVTILKAGBAF 14982  
QY 139 ----- 138  
Db 14983 PLEADVSRPPTMEWSKDGKELEGTAKLEIKIADFSNLNKNKSTRDSDGAYTLTATNP 15042  
QY 139 ----- 138  
Db 15043 GGFAKHIFNVKVLDRPGPEGLAVTEVTSEKCVLSWPPPLDDGGAKIDHVIQKRETSR 15102  
QY 139 ----- 138  
Db 15103 LAWTVASEVQVTKLVTKLLKGNEYIFRVMVNNKYGGEPELESEPVLA VNPYGPDPDPK 15162  
QY 139 ----- 138  
Db 15163 NPEVTTITKDSMVVCGHPDSDGSGSEIINYIVERREDKAGQRWIKCNKKTLLDRLRYKVSGL 15222  
QY 139 ----- 138  
Db 15223 TEGHEYEFRIANAAGISAPSTSPFYKACDVTVPKPGPNPRVLDTSRSSISIAWKNP 15282  
QY 139 ----- 138  
Db 15283 IYDGGSEITGYMVEIALPEDEWQIVTPPAGLKATSYTITGLTENQEKYIRIYAMNSGL 15342  
QY 139 ----- 138  
Db 15343 GEPALVPCTPKAEDRMPLPELADLRKVTIRACCTLRFLVPIKGRPAPEVKNWARDHG 15402  
QY 139 ----- 138

Db 15403 ESIDKASIESTSSYTLIIIVGNVNRFDGKGYILTVENSNGSKSAFVNVVRVLDTPGPPQDLK 15462  
QY 139 ----- 138  
Db 15463 VKEVTKTSVTLTWDPLDGGSKIKNYIIVEKRESTRKAYSTVATNCHKTSWKVDQLQEGC 15522  
QY 139 ----- 138  
Db 15523 SYFVRVLAENEYIGLPAETAESVKASERPLPGKITLMDVTRNSVLSWEKPEHDDGSR 15582  
QY 139 ----- 138  
Db 15583 ILGYIVEMOTKSGDKWATCATVKVTEATITGLIQGEYSFRVSAQNEKGISDPRQLSVPV 15642  
QY 139 ----- 138  
Db 15643 IAKDLVIPAPKLLFNFTFVLAGEDLKVDPFGRPTPAVTHKDNVPLKQTRVNAEST 15702  
QY 139 ----- 138  
Db 15703 ENNSLLTIKDACREDVHYVVKLTNSAGEAIELNVILDKPGPTGPKVMDVETADTSIT 15762  
QY 139 ----- 138  
Db 15763 LSWGPPKYDGGSSINNYIVEKRDSTTTTQWIVSATVARTTIKACRLKTCGEYQFRIAEN 15822  
QY 139 ----- 138  
Db 15823 RYCKSTYLNSEPTVAQPKVPGPCPTPVVTLSSRDSMEVQNEPISDGGSRVIGYHLER 15882  
QY 139 ----- 138  
Db 15883 KERNSILWKLNKTIPIOTKFKTKTGLEGEVEYEFVSAENIVGIGKSKVSECYVARDPC 15942  
QY 139 ----- 138  
Db 15943 DPPGRPEALIVTRNSVTLOWKKPTVDGSKITGYIVEKKELPEGRMMKASFNTIIDTHPE 16002  
QY 139 ----- 138  
Db 16003 VTGLVEDHRYEFVRVARNAGVFSESESTGAITARDEVDPRI SMDPKYKDTIVVHAGE 16062  
QY 139 ----- 138  
Db 16063 SFKVDADYKPIPTIOWIKGQDLSNTARLEIKSTDFATSLSKDVRVDSGNVILRAK 16122  
QY 139 ----- 138  
Db 16123 NVAGERSVTNVKVLDRPGPPGPPVVISGVTAEKCTLAWKPLQDGGSDIINIYIVERRET 16182  
QY 139 ----- 138  
Db 16183 SRLVWTVVDANVQTLSCVKYTKLLEGNEYTFRIMAVNKYGVGEPLSEPVVAKNPFVWPA 16242  
QY 139 ----- 138  
Db 16243 PKAPEVTTVKDSMIVWVERPASDGGSEILGYVLEKRDKEGIRWTRCHKRLIGELRLRT 16302  
QY 139 ----- 138  
Db 16303 GLIENHDEFRVSAENAGLSEPPSPAYQKACDPIYKPGPPNPKVIDITRSSVFLSWS 16362  
QY 139 ----- 138  
Db 16363 KPIYDGGCEIQGYIVEKCDVSGVETMCTPTGINKNTIEVEKLEKHEYNFRICAINKA 16422  
QY 139 ----- 138  
Db 16423 GYGEHADVPPIVEEKELEAPDLDLELRKIINIRAGGSLRFPVPIKGRPTPEVKWGV 16482  
QY 139 ----- 138  
Db 16483 DGEIRDAAIIDVTSFTSLVLDNVNRYDSGKYTLTLENSSGTKSAFVTVRVLDTPSPPVN 16542

QY 139 ----- 138  
Db 16543 LKVEITIKDSVSTWEPPLLDGGSKIKNYIIVEKREATRKSAAVVTNCHNSHKIDQLOE 16602  
QY 139 ----- 138  
Db 16603 GCSYFVRVTAENEYIGLPAQTAADPIKVAEVPQPPGKITVDDVTRNSVLSWTKPEHDGG 16662  
QY 139 ----- 138  
Db 16663 SKIIQYIVEMOAKHSEKWKSECARVKSLOAVITNLTOGEEYLFRRVAVNEKGRSDPRSLAV 16722  
QY 139 ----- 138  
Db 16723 PIVAKDLVIEPDVKPAFSSYSVQVGODLKIEVPISGRPKPTTWTGDKGLPLQKQTRINVT 16782  
QY 139 ----- 138  
Db 16783 DSLDLTTLTSIKETHKDDGGGYGITVANVYGQKTASIEIVTLDKPDPPKGVKFDVSAES 16842  
QY 139 ----- 138  
Db 16843 ITLSWNPPLYTGCCQITNYIVOKRDTTTVWDVVSATVARTTLKVTKLATGTEYQFRI 16902  
QY 139 ----- 138  
Db 16903 ENRYGOSFALES DPIVAQYPIYKEPGPGTFFATAISKDSWVQWHEPVPVNGGSPVIGYHL 16962  
QY 139 ----- 138  
Db 16963 ERKERNILWTKVNKTI IHDTQFKAQNLBEGIEYEFVRYAENIVGVKASKNSECYVARD 17022  
QY 139 ----- 138  
Db 17023 PCDPPTPEPIMVKRNEITLQWTKPVYDGGSMITGYIVEKRDLPGRMMKASFNTVNETQ 17082  
QY 139 ----- 138  
Db 17083 FTVSGLTEDQRYEFVRVIAKNAAGAIKSPSDSTGPIAKDEVELPRISMDPKFRDITVWNA 17142  
QY 139 ----- 138  
Db 17143 GETFRLEADVHGKPLPTIEWLRGDKIEBSARCEIKNTDFKALLIVKDAIRIDGGYILR 17202  
QY 139 ----- 138  
Db 17203 ASNVAGSKSPVNVKVLDRPGPPEGVQVGTSEKSLTWSPLQDGGSDISHYVVEKR 17262  
QY 139 ----- 138  
Db 17263 ETSRLANTVAVSEVVTNSLVTKLLEGNEYVFRIMAVNKYGVGEPLSEAPVLMKNPFVLP 17322  
QY 139 ----- 138  
Db 17323 GPPKSLEVTNIAKDSMTVCWNRPDSDGGSEIIGYIVEKRDGIRWIKCNKRRTDLRLR 17382  
QY 139 ----- 138  
Db 17383 VTGLTEDHEYEFVRVSAENAGVGEPSATVYVYKACDPVKPGPPTNAHIVDTFKNSITLA 17442  
QY 139 ----- 138  
Db 17443 WCKPIYDGGSEILGYVVEICKADEEWOIVTPQTLRVTRFEISKLEHQEYKIRVCALN 17502  
QY 139 ----- 138  
Db 17503 KVGLGEATSVPGTVKPEKLEAPELDLDSELKGIWRAGGSARTHIPFKGRPTPEITWS 17562  
QY 139 ----- 138  
Db 17563 REGEFTDKVTEKGVNTYQLSIDNCDNRDAGKYILKLENSSGSKSAFVTVKVLDTPGPP 17622

QY 139 ----- 138  
Db 17623 QNLAVKEVRKDSAFVWEPIIDGGAKNVVIDKRESTRKAVANVSSKCSKTFKVENL 17682  
QY 139 ----- 138  
Db 17683 TEGAIYFRYMAENEFVGVPVETVDVAKAAEPSPGKVTLLTDVQSOTSASLWKEPEHD 17742  
QY 139 ----- 138  
Db 17743 GGSRLGVVYVEMQPKGTEKHSIVAESKVCNAVVTGLSSGOEYQFRVKAYNEKKS DPRVL 17802  
QY 139 ----- 138  
Db 17803 GVPVIADLFIQPSLKLPPNTYSIQAGEDLKIPIVIGRPPRPNISWVKDGEPLKQTRVN 17862  
QY 139 ----- 138  
Db 17863 VEETATSTVLHIKEGNKDDFGKTYVTATNSAGTATENLSVIVLEKPGPPVGVRFDEVSA 17922  
QY 139 ----- 138  
Db 17923 DFVVISWEPAYTGGCOISNVIYEKKRDTTTHWMSATVARTTIKITKLKTEYQFRI 17982  
QY 139 ----- 138  
Db 17983 FAENRYKSAPLDSKAVIQVPEKEPGPGTPTFTSISKDQMLVQWHEPVNDGGTKIIGY 18042  
QY 139 ----- 138  
Db 18043 HLEQKEKNSILWKLNKPTIQDTFKFTTGLDEGLEEFYKVAENIVIGRPSKVSECFVA 18102  
QY 139 ----- 138  
Db 18103 RDCDPPRPEAIVITRNNTLKWKPAYDGGSKITGYIVEKKDLDPGRWMMKASFNVLE 18162  
QY 139 ----- 138  
Db 18163 TEFTVSLVEDQRYEFRVIAAAGNFSEPSDSSGATTAARDEIDAPNASLDPKYKDVIVV 18222  
QY 139 ----- 138  
Db 18223 HAGETFVLEADIRGKPIPDVVWSKDGKELEETAARMEIKSTIQKTTLLVVKDCIRTDGGQY 18282  
QY 139 ----- 138  
Db 18283 ILKLSNVGGTKSIPITVKVLDLRFPGPEGLKVTGVTAEKCYLAWNPPLODGGANISHYII 18342  
QY 139 ----- 138  
Db 18343 EKRETSRLSWTQVSTEVOALNYKVTLLPGNEYIFRVMVAVNKYIGEPLESGPVTACNPY 18402  
QY 139 ----- 138  
Db 18403 KPPGPPPTPEVSAITKDSMVVTWARPVDDGGTEIGEYILEKRDKEGVRTWKCNKTLTDL 18462  
QY 139 ----- 148  
Db 18463 RLRTVGLTEGHSYEFRAENAAGVGEPSFVYRACDALYPGPPSNPKVTDTSRSSV 18522  
QY 149 ----- 148  
Db 18523 SLAWSKPIYDGGAPKGVVVEVKEAADEWTTCTPTPLGQKQFTVTKLKENTENFRIC 18582  
QY 149 ----- 151  
Db 18583 AINSEGVGEATLPFGSVVAQERIEPPELADLRKVVLRSATLRLFTVIRGRPEPEV 18642  
QY 152 ----- 151  
Db 18643 KWEKAEGILTDRAQIEVTSSTMLVIDNVTRFDSGRYNLTLENNSGSKTAFVNVRLDSP 18702  
QY 152 ----- 151

Db 18703 SAPVNLITREVKKDSVTLWEPPPLIDGGAKITNVIKEKRETRKAVATITNCTKTTFRI 18762  
QY 152 ----- 151  
Db 18763 ENLOQECSSYFRVLASNEYIGLPAETTEPVPKVSEPLPPGRVTLVDVTRNTATIKWEKP 18822  
QY 152 ----- 151  
Db 18823 ESDGSKIITGVVYVEMQTKSEKSWCTQVKTLEATISGLTAGEEYVFRVAANVEKGRSDP 18882  
QY 152 ----- 151  
Db 18883 RQLGVPVIARDIEIKPSVELPFHTFNVKAREQLKIDVPFKGRFOATVNRKDGOTLKETT 18942  
QY 152 ----- 151  
Db 18943 RVNVSSKVTSLSIKEASKEDVGYELCVSNSAGSTVPTITIVLDRPPGPIRIDEV 19002  
QY 152 ----- 151  
Db 19003 SCDSTIISWNPPEYDGGCOISNVIYEKKETTSTTHIVISOAVARTSIKIVRLTGTSEYQF 19062  
QY 152 ----- 151  
Db 19063 RVCAENRYGKSSYSESSAVVAEYFPSPPGPTPKVVHATKSTMLVTWQVPVNDGGSRV 19122  
QY 152 ----- 151  
Db 19123 GYHLEYKERSIILWANKANKILADIATQMKVSGLDEGLMEYRYVAENIAGIKCSKCEPV 19182  
QY 152 PSRD----- 155  
|:|:  
Db 19183 PARDCPPCQPEVTNITRKSLSKSPHYDGGAKITGYIVERRELPGRWLKCNTNI 19242  
QY 156 ----- 155  
Db 19243 QETYFEVTELTEDQRYEFRVFARNAADSVSESTGPIIVKDDVEPRVMDVKFRDVI 19302  
QY 156 ----- 155  
Db 19303 VVKAGEVLKINADIAGRPLPVISWAKDGEIERARTEIISTDNHLLTVKDCIRRTGQ 19362  
QY 156 ----- 155  
Db 19363 YVLTKNVAGTRSVAVNCKVLDKPGPPAGPLEINGLTAECSLSWGRPOEDGGADIDYI 19422  
QY 156 ----- 155  
Db 19423 VEKRETSHLAWTICEGLOMTCKVTLLKNGEYIFRVTGNKYGVGEPLSVAKALDP 19482  
QY 156 ----- 155  
Db 19483 FTVPSPPTSLEITSVTKESMTLWCRSPESDGGSEISGYIITERREKNSLRVVRNKKPVYD 19542  
QY 156 ----- 155  
Db 19543 LRVKSTCLREGCEYEYRYVAENAGLSLPSETSPHIRAEDPVFLPSPPKPIVDSCKTT 19602  
QY 156 ----- 155  
Db 19603 ITIAWVKPLFDGGAPITGYTVEYKKSDTDWKTSIQSLRGTEYITISGLTTGAEYVFRVKS 19662  
QY 156 ----- 155  
Db 19663 VNVKASDPSDSDPQIAKEREEREPLEPDIDSEMRKTLIVKAGASFTVTPFRGPRPVNL 19722  
QY 156 ----- 155  
Db 19723 WSKPDTDLTRAYVDITDSRTSLTIENANRDSCKYTLTIQNVLSAASLTLVVKVLDTPG 19782  
QY 156 ----- 155

Db	19783	PPTNITVQDVTKESAVLSWDVPENDGGAPVKNYHIEKREASKAWSVTNNCNRLSYKVT	19842
QY	156	-----	155
Db	19843	NLOEGAIYFRVSGENFVGIGIPAEKTEGKVKITERPSPPEKLGVTISIKSDSVSLTWLKE	19902
QY	156	-----	155
Db	19903	HDGSRIVHVVEALEKGKNWKAVAKSTHHVYSGLRENEYFRVFAENQAGLSDR	19962
QY	156	-----	155
Db	19963	ELLPLVIKEOLEPPEIDMKNFPSHTVYVRAGSNLKVDIPIISGKPLPKVTLSDRGVPLKA	20022
QY	156	-----	155
Db	20023	TMRNTEITAENLITNLKESVTADAGRYEITAANSGGTTKAFINIVLDRPGPPTGPVVI	20082
QY	156	-----	155
Db	20083	SDITEESVTLKWEPPKYDGGSOVTNYILLKRETSTAVTEVSATVARTMMKVMKLTGEE	20142
QY	156	-----	155
Db	20143	YQFRIKAENRFGISDHIDSACVTVKLPYTTGPPSTPMWTVNTRBSITVGWHEPVSNGS	20202
QY	156	-----	155
Db	20203	AVVGHLEMKDRNSILWOKANKLVTRTHFKVTTISAGLIIEFRVYAENAGVKPSHPS	20262
QY	156	-----	164
Db	20263	EPVLADACEPPRNVRITDISKNSVLSWQOPADFGGSKIITGYIVERDPLDGRWTKASF	20322
QY	165	-----TCL-----	167
Db	20323	TNWTETOITISGLTONSOYEFRRFARNAVSGISNPSEVVGPIITCIDSYGGPVIDLPLEYT	20382
QY	168	-----	167
Db	20383	EVVKYRAGTSVKLRAGISGKPAPTIEWYKDDKELQTNALVCVENTDILASILIKDADRLN	20442
QY	168	-----	167
Db	20443	SGCYELKLRNAGSASATIRVQLDKPGPPGPIEFKTVTAEKITLLWRPPADGGGAKIT	20502
QY	168	-----	167
Db	20503	HYIVEKRETSRVVMSWSEHLECIITTTIKIGNEYIFRVRVKNYKIGIGEPLESDSWA	20562
QY	168	-----	167
Db	20563	KNAFVTPGPGIPEVTKITKNSMTVVWSRPIADGGSDISGYFLEKRDKSLGWFVKLKET	20622
QY	168	-----	167
Db	20623	IRDTRQVKTGLTENSIDYQYRCVAVNAACQGFSEPSFYKAADPIDPPGPAKIRIADST	20682
QY	168	-----	167
Db	20683	KSSITLGWSPKYDGGSAVTGYVVEIROGEEEEWTVTKGEVRTTEVYVSNLKPGVNYI	20742
QY	168	-----	167
Db	20743	FRVSAVNCAGOGEPTEMNEPVQAKDILEAPEIDLVALRTSVIAKAGEDVQLIPFKGRP	20802
QY	168	-----	167
Db	20803	PPTVTRKDEKNLGDARYSIENTDSSLLTIPOVTRNDTGKYILTIENGVGEPKSTVS	20862
QY	168	-----	167
Db	20863	VKVLDPACQKLOVKHVSRGTVTLLWDPPLIDGGSPINNVIEKRDATKRTSVVSHKC	20922

[illegible]

QY 184 ----- 183  
Db 22003 EKCTROIFKVNDLAEGVPIYFRYSANVEYGVGEPIYEMPEPIVATEQAPPRRLDVDTSK 22062  
QY 184 ----- 183  
Db 22063 SSAVLAWLKPDDGSGRITGYLLEMRQKGSDFWEAGHTKQLTFTVERLVEKTEYEFVRK 22122  
QY 184 ----- 183  
Db 22123 AKNDAGYSEPREAFSSVIIKEPOIEPTADLTGITNOLITCKAGSPFTIDVPSGRPAKV 22182  
QY 184 ----- 183  
Db 22183 TWKLEEMRLKETDRVSITTTKDTTLTVKDSMRGDSGRFLTLTENTAGVKTFSTVTVVIG 22242  
QY 184 ----- 186  
QY 184 -----GQP-----  
|:|  
Db 22243 RGPVTPGPIEVSSVSAESCVLWSGEPKGGTEITNYIVEKRESGTTAWQLVNSSVKRTQ 22302  
QY 187 ----- 186  
Db 22303 IKVTHLTWMEYSFRVSSNRFVSKPLSAPIIAEHFPVPSAPTRPEYHVHSANAMSI 22362  
QY 187 ----- 193  
QY 187 -----ENNYKTT-----  
| | | | |  
Db 22363 RWEPEYHDGSKIIGYWEKERTILWVKENKVPCLCNCYKVTGLVEGLEQFRTYALN 22422  
QY 194 ----- 193  
Db 22423 AAGVSKASEASRPIMAQNPVDPAGRPEDVTRSTVLSIWSAPAYDGGSKVVGYYIERKP 22482  
QY 194 ----- 193  
Db 22483 VSEVGDGRWLKCNYYIVSDNFFVTVALSEGDYEFVRVLAKNAAGVTSKSESTGPVTCRD 22542  
QY 194 ---pp-----  
| |  
Db 22543 EYAPPAELDARLHGLVTRAGSDLVDAAVGGKPEPKIITWKGDKELDLCEKVSQYT 22602  
QY 196 ----- 201  
QY 196 -----VLDSG-----  
| | | | |  
Db 22603 GKRAVIFCDRSDSGKTYLTVKNASGKAVSMVKVLDSPGCGKLTAVSRVTQEKCTL 22662  
QY 202 ----- 201  
Db 22663 ANSLPOEDGGAETHYIVERRETSRLNWIVEGECPTLSYVVTRLIKNNEYIFRVRVAVNK 22722  
QY 202 -----SFFL-----  
| | |  
Db 22723 YGPGVPESEPIVARNSTIPSPGPIEVEGTGKEHIIQWTKPESDGGNEISNYLVDKR 22782  
QY 206 ----- 205  
Db 22783 EKSLRWKNDYVYDTRKVTSLMEGCDYQFRVAVNAAGNSBPSEASNFISCREPS 22842  
QY 206 ----- 205  
Db 22843 YTPGPPSAPRVDDTKHSISLAWTKPMYDGGTDIVGYVLEMKEKTDQWVRVHTNATIRN 22902  
QY 206 ----- 205  
Db 22903 TEFTVPLKMGQYSFRAAVNVKMGSESAETIEPVERIEIPDLELADDLKKTVTIR 22962  
QY 206 ----- 213  
QY 206 -----YSKLTVDK-----  
| | | | |  
Db 22963 AGASRLWVSGRPPPVITWSQGDILASRAIIDTTESYSLIIVDKVNRIDAGKYTIEA 23022  
QY 214 ----- 213  
Db 23023 ENOSGKKSATVLVKVYDTPGCPBSVKVKEVSRDSVTITWEIPTDGGAPVNNVIVEKREA 23082  
QY 214 ----- 213

Db 23083 AMRAFKVTTKCKSKTLYRISGLVEGTMYFRVLPENIYGICEPCETSDAVLVSEVPLVPA 23142  
QY 214 ----- 213  
Db 23143 KLEVVDVTKSTVTLAWEKPLYDGGSRITGYVLEACKAGTERMMKVVTLPKPTVLEHTVTSL 23202  
QY 214 ----- 213  
Db 23203 NEGEQYLFRIRAONEKGVSEPRETVTAVTVQDLRVLPFTIDLSTMPQKT IHVPAGRPVELV 23262  
QY 214 ----- 213  
Db 23263 IPIAGRPPPAASWFFAGSKLRESERVIVETHTKVAKLTIRETTIROTGEYTLLEKKNVTGT 23322  
QY 214 ----- 213  
Db 23323 TSETIKVIIIDKPPPTGPIKIDEIDATISITISWEPPELDGGAPLSGYVVVEQDAHRPGW 23382  
QY 214 ----- 213  
Db 23383 LPVSESVTRSTFKFTRLTEGNEVYFRVAATNRFGIGSYLQSEVIECRSSIRIPGPETLQ 23442  
QY 214 ----- 213  
Db 23443 IPDVSRDGMTLTWYPPEDDGGSOVTGYIVERKEVRADRVVRVNVKVPVTMYRSTGLTEG 23502  
QY 214 ----- 213  
Db 23503 LEYEHRTAINARGSGKPSRSPKPIVAMPDIAPGPKQNPVRVTDTTTTSVSLAWSPEDE 23562  
QY 214 ----- 213  
Db 23563 GSKVTGYLIEMQVQDHEWTKCNTPTKIREYTLHLPOGAERYRVLACNAGGPEPA 23622  
QY 214 -----SRMOQG-----  
|:|:|  
Db 23623 EYVGTVKVTEMLEYPDYELDERYQEGIFVRQGGVIRLTIPIKGPFPICKWKEGQDISK 23682  
QY 220 ----- 219  
Db 23683 RAMIATSETHTELVIKEADRGDSGTVDLVLENKCGKKAIVIKRVITGSPNSPEGPLEYDD 23742  
QY 220 ----- 219  
Db 23743 IQVRSVRSWRPPADGGADILGYILERREVPKAAWTTIDSRVRCYTSLVVKGLKENVEYH 23802  
QY 220 ----- 219  
Db 23803 FRVSAENQFGISKPLKSEEPVTEKTLNPPPEPPNPPELVDVTKSSVLSWSRPKDDGGS 23862  
QY 220 ----- 219  
Db 23863 RVTGYIERKETSTDKWRHNTQITTTMYTVTGLVPAEYQFRIIAQNADVGLSETSPAS 23922  
QY 220 ----- 219  
Db 23923 EPVVCCKDPDKPSQPGELEILSISKDSVTLOWEKEPCDGGKEILGYWVEYRQSGDSANWK 23982  
QY 220 ----- 219  
Db 23983 SNKERIKDKQFTIGGLLEATEYEFVFAENETGLSRPRRTAMSIKTTLTSGEAPGIRKEM 24042  
QY 220 ----- 219  
Db 24043 KDVTTKLGEAAALSCQIVGRPLDIXWRFGKELIOSRKYKMSDGRTHLTVMTEEQED 24102  
QY 220 ----- 219  
Db 24103 EGVYTCIATNEVEVETSSKLLLOATPQPHPGYPLKEKYGVAGVSTLRLHVMYIGRPVPA 24162  
QY 220 ----- 222  
| | |  
| | |

Db 24163 MTWFHQKLLQNSENIITIENTEHTYHLVKNVQRKTHAGKYKVLNSVFGTVDAILDVEI 24222  
QY 223 ----- 222  
Db 24223 QDKPDKPTGPIVIEALLKNSAVISKHPADGGSWITNVVVEKCEAKEGAEWLVSSAIS 24282  
QY 223 ----- 222  
Db 24283 VTTCRIVNLTENAGYFRVSAQNTFGISDPLEVSSWIIKSPFKPGAPGKPTITAVTKD 24342  
QY 223 SCSV----- 226  
|| |  
Db 24343 SCVWAKPPASDGGAKIRNYYLEKREKKONKWISVTTEIRETVFSVKNLIEGLEYEFRV 24402  
QY 227 ----- 226  
Db 24403 KCENLGGESEWSEISEPTTPKSDVPIQAPHKEELRNLRVYQSNATLVCKVTGHPKPIV 24462  
QY 227 ----- 226  
Db 24463 KWYRQKEIIDGLKYRIQEFKGYHQLIIASVTDDATVYQVRATNOGGSVSGTASLEV 24522  
QY 227 ----- 226  
Db 24523 EYPAKIHLPKLEGMGAVHALRGEVSIKIPFGKPDPIITWQGGDLIDNNGHYQIVT 24582  
QY 227 ----- 226  
Db 24583 RSFTSLVPNGVERKDAGFYVVCANRFGIDQKTVELDVADVPDPGRGVKVSVDVSRDSVN 24642  
QY 227 ----- 226  
Db 24643 LTWTEPASDGGSKITNYIVEKATTAERWLVRGQARETRYVINLFGKTSYQFRVIAENK 24702  
QY 227 ----- 226  
Db 24703 FGLSKPSEPTITTKEDKTRAMNYDEEDETREVSMTKASHSTKELYERYMIAEDLGR 24762  
QY 227 ----- 230  
MHEA-----  
|| |  
Db 24763 GEFGIVHRCVETSSKKTYMAFVKVKGTDQVLVKEISILNIAHRNLIHLHESFESMEE 24822  
QY 231 ----- LNNH----- 234  
|| |  
Db 24823 LVMIFISGLDIFERINTSAFELNEREIVSVHQVCEALQFLSHNIGHFDIRPENIY 24882  
QY 235 ----- 234  
Db 24883 QTRRSSTIKIIEFGQAROLKPGDNFRLLFTAPEYYAPEVHOHDVVSTATDMWSLGLTVYV 24942  
QY 235 -----YT----- 236  
|| |  
Db 24943 LLSGINPFLAETNQOIENINNAEYTFDEAFKEISIEAMDFDRLLVKERSRMTASEA 25002  
QY 237 -----QKSLSLS- 243  
|| |  
Db 25003 LQHPWLKOKIERVSTKVRTLKHRRYHTLIKDLNMVMSAARISCGGAIRSQRGSVAK 25062  
QY 244 ----- 243  
Db 25063 VKVASIEIGPVSGQIMHVAEGEGHVYKVKIENYDOSTQVTWYFGVQLENSEKEYEITY 25122  
QY 244 ----- 243  
Db 25123 EDGVAILVVKDITKLDDGTYRCKVNDYGEDSSYAELFVGKGVREYDYCYCRTMKIKRR 25182  
QY 244 ----- 243  
Db 25183 TDTMLLERPPEFTLPLYNKTYAYGENVRFGVITVHPHPTWTWYKSCQIKPGDNCKY 25242  
QY 244 ----- 243  
Db 25243 TFSDKGLYQLTINSVTDDDAEYTVVARNKYGEDSCAKLTVTLHPPPTDSTLRPMFKR 25302

QY 244 ----- 243  
Db 25303 LLANACQEGOSVCFEIRVSGIPPTLKWEXDQGLSLGPNIEIIHEGLDYVALHIRDTL 25362  
QY 244 ----- 243  
Db 25363 PEDTGYRVTTATNTAGSTSCQAHLQVERLRYKQBEFKSKEEHERHVQKQIDKTLRMAEIL 25422  
QY 244 ----- 243  
Db 25423 SGTESVPLTQVAKREALREAAVLYKPAVSTKTVKGEFRLEIEEKKERKLMPYDVPEPRK 25482  
QY 244 -----PGK 246  
|| |  
Db 25483 YKQTTIEEDQRIKQFVPMDSMDMKWYKKIRDOYEMPGK 25518

Search completed: April 21, 2003, 10:50:13  
Job time : 99.1765 secs



GenCore version 5.1.4.p5.4578  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: April 21, 2003, 10:42:21 ; Search time 31.0588 Seconds  
(without alignments)  
1132.630 Million cell updates/sec

Title: 7LINK7LINK2  
Perfect score: 1430  
Sequence: 1 DNLKAFYDKVAEKLEAFDW.....MHEALHHVYTKSLSLSPGK 264

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.0

Searched: 908470 seqs, 133250620 residues

Total number of hits satisfying chosen parameters: 908470

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : A\_Geneseq\_101002.\*

- 1: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1980.DAT.\*
- 2: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1981.DAT.\*
- 3: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1982.DAT.\*
- 4: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1983.DAT.\*
- 5: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1984.DAT.\*
- 6: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1985.DAT.\*
- 7: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1986.DAT.\*
- 8: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1987.DAT.\*
- 9: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1988.DAT.\*
- 10: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1989.DAT.\*
- 11: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1990.DAT.\*
- 12: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1991.DAT.\*
- 13: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1992.DAT.\*
- 14: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1993.DAT.\*
- 15: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1994.DAT.\*
- 16: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1995.DAT.\*
- 17: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1996.DAT.\*
- 18: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1997.DAT.\*
- 19: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1998.DAT.\*
- 20: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1999.DAT.\*
- 21: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA2000.DAT.\*
- 22: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA2001.DAT.\*
- 23: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA2002.DAT.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	ID	Description
1	1277	89.3	683	21	AA196781 Ephrin-B1-Ephrin-B
2	1277	89.3	754	21	AA111691 Human secreted K10
3	1277	89.3	787	21	AA111693 Human secreted K10
4	1271	88.9	1218	19	AA170539 Integrin alpha-4 c
5	1271	88.9	1367	19	AA170542 Integrin alpha-2 c
6	1268	88.7	594	22	AAU09817 Murine FGFR-L extr
7	1268	88.7	951	20	AAW70798 Human gp130-C-gamm
8	1268	88.7	951	21	AA192186 Human gp130-C-gamm
9	1268	88.7	961	21	AA192187 Human gp130-J-C-ga
10	1268	88.7	1158	21	AA192205 Fusion polypeptide

11	1268	88.7	1168	21	AA192204 Fusion polypeptide
12	1265	88.5	475	22	AAG63640 Amino acid sequenc
13	1265	88.5	690	21	AA192195 Human IL-6P-alpha-
14	1264	88.4	473	22	AAG64473 Human type antihum
15	1264	88.4	473	22	AAG64475 Human type antihum
16	1263	88.3	567	23	AAE13733 Human zalphalir/ig
17	1263	88.3	859	20	AAW70796 Human gp130-Fc-His
18	1263	88.3	859	21	AA192184 Human gp130-Fc-His
19	1263	88.3	963	19	AAW70540 Integrin beta-1 ch
20	1262	88.3	583	22	AA1983156 Ganglioside GM2 an
21	1261	88.2	445	20	AA124153 Bovine LOX-1 extra
22	1261	88.2	471	21	AA1945030 HUMAN OCR10-Fc fus
23	1261	88.2	473	22	AAG64469 Human type antihum
24	1261	88.2	473	22	AAG64471 Human type antihum
25	1261	88.2	476	18	AAW01822 Primatised anti-hu
26	1261	88.2	476	19	AAW63765 Macaque primatised
27	1261	88.2	476	23	AAU11646 Protein sequence o
28	1261	88.2	628	21	AA115179 GFRA1pha3-IgG fusi
29	1260	88.1	399	21	AA170867 Human interferon-b
30	1260	88.1	449	14	AA173339 Completely humanis
31	1260	88.1	449	19	AAW49816 Amino acid sequenc
32	1260	88.1	497	21	AA197172 Human FGF-R1 Extra
33	1260	88.1	525	21	AA197171 Human FGF-R1 Extra
34	1260	88.1	622	21	AA197170 Human FGF-R1 Extra
35	1260	88.1	633	21	AA184965 Amino acid sequenc
36	1259	88.0	462	21	AA129408 Human monoclonal a
37	1259	88.0	475	18	AAW11639 Human anti-RSV mon
38	1259	88.0	534	13	AA126531 Sequence Of CD4-Ig
39	1259	88.0	547	22	AA185279 Human IL-20RA-Ig g
40	1259	88.0	547	23	AA1852710 Interleukin-20RB/I
41	1259	88.0	547	23	AA123362 Human IL-20RA EC d
42	1259	88.0	571	22	AAU04065 Human IL-20RA/immu
43	1259	88.0	571	22	AA185278 Human IL-20RA-Ig g
44	1259	88.0	571	23	AA1852709 Interleukin-20RB/I
45	1259	88.0	571	23	AA123359 Human IL-20RA extr

ALIGNMENTS

RESULT 1  
AA196781  
ID AA196781 standard; Protein; 683 AA.

XX AA196781;

AC AA196781;

XX 26-SEP-2000 (first entry)

DT Ephrin-B1-Ephrin-B1-FC fusion protein.

XX Ephrin-B1; ELK receptor; ligand; dimer; Fc domain; fusion protein;

DE Ephi-6 antagonist; neurological.

KW Chimeric - Homo sapiens.

OS Chimeric - Synthetic.

XX Key Location/Qualifiers

FT Peptide 1..29

FT /label= Signal\_peptide

FT /note= "derived from ELK-L Ectodomain 1"

FT 30..237

FT /label= ELK-L\_ectodomain\_1

FT 238..240

FT /label= Bridging\_peptide

FT 241..448

FT /label= ELK-L\_Ectodomain\_2

FT 449..451

FT /label= Bridging\_peptide

FT 452..683

FT /label= Human\_IgG1\_Fc\_region

XX WO200037642-A1.

PD 29-JUN-2000.  
XX 23-DEC-1999; 99WO-US30900.  
XX 23-DEC-1998; 98US-0113387.  
XX (REGE-) REGENERON PHARM INC.  
XX Davis SJ, Gale NW, Yancopoulos GD, Stahl N;  
XX WPI; 2000-442670/38.  
XX N-PSDB; AAA51345.  
XX Polynucleotide encoding a fusion polypeptide, useful for promoting  
XX differential function and influencing phenotype, comprises two subunits  
XX containing at least one copy of the receptor binding domain of a ligand  
XX Example 12; Fig 14A-E; 97pp; English.  
XX Production of homogenous forms of clustered ligands is broadly applicable  
XX to improve the affinity and/or increase the activity of a ligand as  
XX compared to the native form of the ligand. Ephrin fusion proteins have  
XX been constructed, which may be useful for treating neurological  
XX disorders. The ephrin fusion proteins are preferably capable of binding  
XX to Etk receptor and are especially Etk-6 antagonists. The fusion proteins  
XX were constructed after it was demonstrated that similar improved  
XX activities could be achieved using Tie-2 receptor ligands.  
XX Angiopoietin-1 (Ang-1) is a Tie-2 receptor ligand and is an agonist for  
XX Tie-2, whereas angiopoietin-2 (Ang-2) is a naturally occurring antagonist  
XX of the Tie-2 receptor. The fibrinogen domains (FD) of Ang-1 and Ang-2 are  
XX the receptor-binding domains and dimerized versions, of e.g. Ang-1-FD-Fc  
XX (Ang-1 fibrinogen domain fused to an Fc domain), can bind to the Tie-2  
XX receptor with much higher affinity than monomeric Ang-1-FD (dimerization  
XX occurs between the Fc components of adjacent molecules). However,  
XX Ang-1-FD-Fc is not able to induce phosphorylation (activate) the Tie-2  
XX receptor on endothelial cells unless it is further clustered with goat  
XX anti-human Fc antibodies. The novel fusion proteins, mutant versions of  
XX Ang-1-FD and Ang-2-FD, were designed that were intrinsically more highly  
XX clustered. Tie-2 agonist fusion proteins may be used as haematopoietic  
XX factors. Tie-2 receptor antagonist fusion proteins may be used to  
XX diagnose or treat, e.g. myeloproliferative or other proliferative  
XX disorders of blood forming organs, e.g. thrombocythemias, polycythemia  
XX and leukemias.  
XX SQ Sequence 683 AA;  
Query Match 89.3%; Score 1277; DB 21; Length 683;  
Best Local Similarity 36.5%; Pred. No. 2.2e-35;  
Matches 245; Conservative 6; Mismatches 12; Indels 409; Gaps 5;  
QY 2 WLKA-----FVSKVAEL-----14  
DB 12 WLVAWVWALCLATPLAKNLEVPVSSNLNPKFLSGGLVIVPKIGDKLDIICPRAEAGR 71  
QY 15 -----14  
DB 72 PYEYKLYLVRPEQAACSTVLDPNVLVTCNRPQEPIRETIKQFSPNMGLEFKKHD 131  
QY 15 -----KEAPDLK- 22  
DB 132 YYITSTNGSLEGNREGGVCRTTKIIMKVQDPNAVTPPEQLTTSRPSKEADNTVKM 191  
QY 23 -----22  
DB 192 ATQAPGSRGLSDGSKHETVNOEKSFGPGASGGSGDPDGFNSKPGKNLEPVSSSL 251  
QY 23 -----AFDKVAEL-----32  
DB 252 NPKFLSGGLVIVPKIGDKLDIICPRAEAGRPEYKLYLVRPEQAACSTVLDPNVLV 311  
QY 33 -----32  
DB 312 CNRPEQEIIRTIKQFSPNMGLEFKKHDYYITSTNGSLEGNREGGVCRTTKMI 371

QY 33 -----KEA-----35  
DB 372 IMKVQDPNAVTPPEQLTTSRPSKEADNTVKMATQAPGSRGLSDGSKHETVNOEKS 431  
QY 36 -----FMDKTHTCPPCPAPELILGGPSVFLFPPKPKDITLMISR 72  
DB 432 GASGGSGDDPDGFNSKPGEPKSCDKTHTCPCPAPELILGGPSVFLFPPKPKDITLMISR 491  
QY 73 TPEVTCVVVDVSHEDPEVKFNWYVDGVEVHNATKPRDEQYNSTYRVVSVLTVLDHDLN 132  
DB 492 TPEVTCVVVDVSHEDPEVKFNWYVDGVEVHNATKPRDEQYNSTYRVVSVLTVLDHDLN 551  
QY 133 GKYEKCKVSNKALPAPIEKTISKAKGQPREPQVYTLPPSRDELTKNQVSLTCLVKGFYPS 192  
DB 552 GKYEKCKVSNKALPAPIEKTISKAKGQPREPQVYTLPPSRDELTKNQVSLTCLVKGFYPS 611  
QY 193 DIAVEWESNGQPENNYKTTPPVLDSDGSFFLYSKLTVDKSRMQQGNVFSCSVMHEALHNH 252  
DB 612 DIAVEWESNGQPENNYKTTPPVLDSDGSFFLYSKLTVDKSRMQQGNVFSCSVMHEALHNH 671  
QY 253 YTKSLSLSPGK 264  
DB 672 YTKSLSLSPGK 683  
RESULT 2  
AAB11691  
ID AAB11691 standard; Protein; 754 AA.  
XX AAB11691;  
XX 23-OCT-2000 (first entry)  
XX Human secreted Klotho/IgG1 hinge-CH2-CH3 chimeric protein.  
XX Human; secreted Klotho/IgG1 heavy chain constant region chimeric protein;  
XX immunoglobulin G1; gene therapy; antibody; antiarthritic; nephrotropic;  
XX kidney disease; oedema; dropsy; arthritis.  
XX Homo sapiens.  
XX Synthetic.  
XX WO200027885-A1.  
XX 18-MAY-2000.  
XX 05-NOV-1999; 99WO-JP06152.  
XX 05-NOV-1998; 98JP-0314153.  
XX (KYOW ) KYOWA HAKKO KOGYO KK.  
XX Hanai N, Nakamura K, Kato Y, Nabeshima Y;  
XX WPI; 2000-376499/32.  
XX N-PSDB; AAA61542, AAA61544.  
XX Chimeric polypeptide containing klotho protein bound to an  
XX immunoglobulin for treatment and prevention of kidney disease, dropsy  
XX and arthritis  
XX Claim 9; Page 59-63; 94pp; Japanese.  
XX This sequence represents a chimeric protein comprising mature  
XX human secreted Klotho protein and the heavy chain constant region  
XX (hinge-CH2-CH3) of human IgG1 (immunoglobulin G1). The invention  
XX relates to such chimeric proteins comprising the human Klotho protein  
XX fused to a human immunoglobulin, and to DNA sequences encoding the chimeric  
XX proteins of the invention, to gene therapy vectors comprising DNA encoding the chimeric  
XX proteins of the invention, to detection and assay of molecules which  
XX interact with Klotho, to antibodies raised against the chimeric protein  
XX and to a diagnostic method for Klotho-related disorders using the

CC antibodies. The chimeric proteins of the invention have antiarthritic and  
CC nephrotropic activity, and may be used for the treatment, prevention and  
CC diagnosis of disorders with which klotho is associated, including kidney  
CC disease, oedema (dropsy) and arthritis.  
XX  
SQ Sequence 754 AA;

Query Match 89.3%; Score 1277; DB 21; Length 754;  
Best Local Similarity 38.0%; Pred. No. 2.8e-35;  
Matches 241; Conservative 10; Mismatches 12; Indels 372; Gaps 5;

QY 2 WLKA-----FYDKVAEKLKE----- 16  
| : : : : :  
Db 120 WARVLPNGSAGVPNREGRLRYRLLERLRELGVQPVVTLYHWDLPQRLQDAYGWNANRAL 179  
QY 17 ----- 16  
Db 180 ADHFRDYAELCFRHFQGVQVYITIDNPYVVAWHGYATGRAPGIRGSPRLGLYVAHNLL 239  
QY 17 ----- 16  
Db 240 LAHAKVWHLYNTSFRPTQGGQVSTALSSHWINPRMTDHSIKECKSLDFVLGWFAPVF 299  
QY 17 ----- 16  
Db 300 IDGYPESMKNLSSILPDFTESEKFKIGTADFFALCFGLTSLFQLLDPHMKFRQLESP 359  
QY 17 ----- 16  
Db 360 NLQQLSWIDLEFNHPQIFIVENGWVSGTTKRDDAKYMYLKKFIMETLKAIKLGDGV 419  
QY 17 -----FYDKVAEK----- 31  
| : : : : :  
Db 420 IGYTANSLMDGFEWHRGYSIRRGIFYDFLSQDKMLLPKSSALFYQKLIKNGFPPLPEN 479  
QY 32 --LKEAF-----MDKTHTCPPCAP 49  
| : : : : :  
Db 480 QLEGTFPCDFAWGVVDNYIQVSQLTKPISSLTKPYHSGSGGEPKSCDKTHTCPPCAP 539  
QY 50 ELGGPSVFLFPKPKDLMISRTPEVTCVVDVSHEDPEVKFNWYDGVGVHNAKTNR 109  
| : : : : :  
Db 540 ELGGPSVFLFPKPKDLMISRTPEVTCVVDVSHEDPEVKFNWYDGVGVHNAKTNR 599  
QY 110 EEQNSTYRVSVLTVLHODWLNKKEYCKVSNKALPAPIEKTISKAKGPREPQVYTL 169  
| : : : : :  
Db 600 EEQNSTYRVSVLTVLHODWLNKKEYCKVSNKALPAPIEKTISKAKGPREPQVYTL 659  
QY 170 PSRDELTKNOVSLTCLVKGFYPSDIAVEWESNGOPENNYKTTTPVLDSDGSFFLYSKLT 229  
| : : : : :  
Db 660 PSRDELTKNOVSLTCLVKGFYPSDIAVEWESNGOPENNYKTTTPVLDSDGSFFLYSKLT 719  
QY 230 DKSRWQGNVFCVSMHEALHNHYTOKLSLSPGK 264  
| : : : : :  
Db 720 DKSRWQGNVFCVSMHEALHNHYTOKLSLSPGK 754

RESULT 3  
AAB11693

ID AAB11693 standard; Protein; 787 AA.

AC AAB11693;

XX 23-OCT-2000 (first entry)

XX Human secreted Klotho precursor/IgG1 hinge-CH2-CH3 chimeric protein.

XX Human; secreted Klotho/IgG1 heavy chain constant region chimeric protein;  
KW immunoglobulin G1; gene therapy; antibody; antiarthritic; nephrotropic;  
KW kidney disease; oedema; dropsy; arthritis.

OS Homo sapiens.

OS Synthetic.

XX

FH Key Location/Qualifiers  
FT Peptide 1..33  
FT Protein /note= "Signal peptide"  
FT Protein 34..787  
FT Protein /note= "Human secreted Klotho/IgG1 heavy chain constant  
FT Protein region chimeric protein (AAB11691)"  
XX WO200027885-A1.  
PN 18-MAY-2000.  
XX  
XX 05-NOV-1999; 99WO-JP06152.  
PF  
XX 05-NOV-1998; 98JP-0314153.  
PR  
XX (KYOW ) KYOWA HAKKO KOGYO KK.  
PA  
XX Hanai N, Nakamura K, Kato Y, Nabeshima Y;  
PI WPI: 2000-376499/32.  
XX N-PSDB; AAA61544.  
DR  
XX Chimeric polypeptide containing klotho protein bound to an  
PT immunoglobulin for treatment and prevention of kidney disease, dropsy  
PT and arthritis  
XX  
PS Disclosure; Page 73-79; 94pp; Japanese.  
XX  
CC This sequence represents a chimeric protein comprising human secreted  
CC Klotho protein precursor and the heavy chain constant region  
CC (hinge-CH2-CH3) of human IgG1 (immunoglobulin G1). The invention  
CC relates to such chimeric proteins comprising the human Klotho protein  
CC fused to a human immunoglobulin, and to DNA sequences encoding them. It  
CC also relates to gene therapy vectors comprising DNA encoding the chimeric  
CC proteins of the invention, to detection and assay of molecules which  
CC interact with Klotho, to antibodies raised against the chimeric protein  
CC and to a diagnostic method for Klotho-related disorders using the  
CC antibodies. The chimeric proteins of the invention have antiarthritic and  
CC nephrotropic activity, and may be used for the treatment, prevention and  
CC diagnosis of disorders with which Klotho is associated, including kidney  
CC disease, oedema (dropsy) and arthritis.  
XX  
SQ Sequence 787 AA;  
Query Match 89.3%; Score 1277; DB 21; Length 787;  
Best Local Similarity 38.0%; Pred. No. 3e-35;  
Matches 241; Conservative 10; Mismatches 12; Indels 372; Gaps 5;  
QY 2 WLKA-----FYDKVAEKLKE----- 16  
| : : : : :  
Db 153 WARVLPNGSAGVPNREGRLRYRLLERLRELGVQPVVTLYHWDLPQRLQDAYGWNANRAL 212  
QY 17 ----- 16  
Db 213 ADHFRDYAELCFRHFQGVQVYITIDNPYVVAWHGYATGRAPGIRGSPRLGLYVAHNLL 272  
QY 17 ----- 16  
Db 273 LAHAKVWHLYNTSFRPTQGGQVSTALSSHWINPRMTDHSIKECKSLDFVLGWFAPVF 332  
QY 17 ----- 16  
Db 333 IDGYPESMKNLSSILPDFTESEKFKIGTADFFALCFGLTSLFQLLDPHMKFRQLESP 392  
QY 17 ----- 16  
Db 393 NLQQLSWIDLEFNHPQIFIVENGWVSGTTKRDDAKYMYLKKFIMETLKAIKLGDGV 452  
QY 17 -----AFDWLKA-----FYDKVAEK----- 31  
| : : : : :  
Db 453 IGYTANSLMDGFEWHRGYSIRRGIFYDFLSQDKMLLPKSSALFYQKLIKNGFPPLPEN 512  
QY 32 --LKEAF-----MDKTHTCPPCAP 49

Db 513 QPLEGTFPCDFAWGVNDVNIQVSQLTKPISSLTQPKYHSGSGGEPKSCDKTHTCPCPAP 572  
QY 50 ELGGSPVFLFPKPKDMLISRTPEVTCVVVDVSHEDPEVKENWYVDGVEVHNATKPR 109  
Db 573 ELGGSPVFLFPKPKDMLISRTPEVTCVVVDVSHEDPEVKENWYVDGVEVHNATKPR 632  
QY 110 EEQNSTYRVVSVLTLDHDLNGKEYCKVSNKALPAPIETKISKAKGQPREPQVYTLTP 169  
Db 633 EEQNSTYRVVSVLTLDHDLNGKEYCKVSNKALPAPIETKISKAKGQPREPQVYTLTP 692  
QY 170 PSDELTKNQVSTCLVKGYFSPDSIAVEWESNGQPENNYKTTPPVLDSDGSFPLYSKLTAV 229  
Db 693 PSDELTKNQVSTCLVKGYFSPDSIAVEWESNGQPENNYKTTPPVLDSDGSFPLYSKLTAV 752  
QY 230 DKSRWQGNVFCFSVMHEALHNHYTOKLSLSLSPGK 264  
Db 753 DKSRWQGNVFCFSVMHEALHNHYTOKLSLSLSPGK 787  
RESULT 4  
AAW70539  
ID AAW70539 standard; Protein; 1218 AA.  
XX AC AAW70539;  
XX DT 26-JAN-1999 (first entry)  
XX DE Integrin alpha-4 chain.  
XX KW Integrin; alpha-4 chain; immunoglobulin; chimeric; heterodimer complex;  
XX KW inhibitor; binding; ligand; blood platelet; hemostatic; diagnostic agent;  
XX KW human.  
XX OS Homo sapiens.  
XX FH Key  
FT Peptide 1..39 Location/Qualifiers  
FT /note= "signal peptide"  
FT Protein 40..1218  
FT /note= "mature protein"  
XX W09832771-A1.  
XX 30-JUL-1998.  
XX 29-JAN-1998; 98WO-JP00370.  
XX 29-AUG-1997; 97JP-0234544.  
XX 29-JAN-1997; 97JP-0015118.  
XX (TORA ) TORAY IND INC.  
XX PA Kainoh M, Tanaka T;  
XX PI WPI; 1998-427881/36.  
XX DR N-PSDB; AAV33772.  
XX Integrin-immunoglobulin chimeric protein heterodimer complexes as  
PT platelet substitutes - contain the alpha and beta integrin chains  
PT associated in stable state and bind to extracellular matrix in the  
PT presence of plasma components  
XX Claim 7; Pages 42-50; 87pp; Japanese.  
XX This represents an integrin alpha-4 chain. The invention provides  
CC integrin-immunoglobulin chimeric protein heterodimer complexes that  
CC comprise an integrin alpha or beta chain associated with an  
CC immunoglobulin light or heavy chain. These chimeric proteins form  
CC heterodimer complexes, in particular with a chimeric protein containing  
CC an integrin alpha chain and an immunoglobulin chain with a chimeric  
CC protein containing an integrin beta chain and an immunoglobulin chain;  
CC the immunoglobulin chain in each case may be a heavy chain, or one of the

CC two may be a light chain. The integrin alpha chain is preferably alpha 4  
CC or alpha 2 and the integrin beta chain is preferably beta 1. Animal cells  
CC transformed with vectors containing the DNA coding for the above chimeric  
CC proteins can be used in the preparation of the chimeric proteins and  
CC their heterodimer complexes. The heterodimer complexes, which are useful  
CC for testing potential promoters and inhibitors of the binding of  
CC integrins to their ligands, function as blood platelet substitutes and  
CC hemostatics and as diagnostic agents.  
XX Sequence 1218 AA;  
SQ Query Match 88.9%; Score 1271; DB 19; Length 1218;  
Best Local Similarity 24.9%; Pred. No. 1.3e-34;  
Matches 243; Conservative 4; Mismatches 14; Indels 716; Gaps 4;  
QY 4 KAFYDK----- 9  
Db 242 KAFDKQNVKFGSYLGYSVGAGHFRSQHTTEVVGAPQHEQIGKAYIFSIDEKELNHLH 301  
QY 10 ----- 9  
Db 302 EMKGKLGSYFGASCAVDLNADGFSDDLVLGAPMQSTIREGRVFVYINSGGAVNAME 361  
QY 10 ----- 9  
Db 362 TNLVGSKYAARFGESIVNLGIDIDNGFEDVAIGAPQEDDLQGAIIYINGRADGISSTFS 421  
QY 10 ----- 9  
Db 422 QRIEGLQISKLSMFQSGISGQIDADNNGYVDVAVGAFRSDSAVLLTRPVVIVDASLSH 481  
QY 10 ----- 9  
Db 482 PESVNRKFCDCVENGWPSVICIDLTLCSYKGVPCYIVLVFNWMSLDVNRKAESPPEYF 541  
QY 10 ----- 9  
Db 542 SSGTSDVITGSTQVSSREANCRTHOAFMRKDVDRDILTPQIEAAVHLGPHVSKRSTEE 601  
QY 10 ----- 9  
Db 602 FPPLPILQKKKEKDIMKKTINFARFCAHENSADLQVSAKIGFLAPHENKYLAVGSMK 661  
QY 10 ----- 9  
Db 662 TMLNVSFLNAGDDAYETTLHVKLPGVLYFIKILEEEKQINCEVTDNSGVQLDCSIGV 721  
QY 10 ----- 9  
Db 722 IYVDHLRIDISFLLDVSSLSRAEEDLSITVHATCENEEMDNLKHRSVRTVAIPLKYEVK 781  
QY 10 ----- 9  
Db 782 LTVHGFVNPTSFYGSNDENEPETCMVEKMNLTFFHINTGNSMAPNVSVIWPVNSFSPQ 841  
QY 10 -----VAEKLKEAFOWLKA---FYDKVAEKL----- 32  
Db 842 TDKLFNLVDVQTGTGECHEFNQVRCALQEQKSAQMTLQAGIVRFLSKTDKRLLYCIKADP 901  
QY 33 -----KEAFM----- 37  
Db 902 HCLFLNLCNFGKMSGKASVHIQLEGRPSILEMDETSAKFEIRATGFPENPRVIELNK 961  
QY 38 -----DKTHCTPCPAPPELLGGSPVFLFPKPKDT 67  
Db 962 DENVAHVLEGLUHQRPKRYFTDPEEPKSCDKTHTCPCPAPPELLGGSPVFLFPKPKDT 1021  
QY 68 LMISRTPEVTCVVVDVSHEDPEVKFNWYVDGVEVHNATKPREEOYNSTYRVVSVLTVLH 127  
Db 1022 LMISRTPEVTCVVVDVSHEDPEVKFNWYVDGVEVHNATKPREEOYNSTYRVVSVLTVLH 1081  
QY 128 QDWLNGKEYCKVSNKALPAPIETKISKAKGQPREPQVYTLPPSRDELTKNQVSLTCLVK 187  
|||||

Db 1082 QDLNGKEYCKVSNKALPAIEKTIISKARQPREPOVYTLPPSRDELTKNOVSLTCLVK 1141

QY 188 GFYPSDIAVWESNGQPPENNYKTTTPVLDSDGSGFFLYSKLTVDKSRWQOGNVFSCSVME 247  
|||||

Db 1142 GFYPSDIAVWESNGQPPENNYKTTTPVLDSDGSGFFLYSKLTVDKSRWQOGNVFSCSVME 1201

QY 248 ALHNHYTKQSLSLSPGK 264  
|||||

Db 1202 ALHNHYTKQSLSLSPGK 1218

RESULT 5

AAW70542

ID AAW70542 standard; Protein; 1367 AA.

XX

AC AAW70542;

XX

DT 26-JAN-1999 (first entry)

XX

DE Integrin alpha-2 chain.

XX

KW Integrin; alpha-4 chain; immunoglobulin; chimeric; heterodimer complex;

KW inhibitor; binding; ligand; blood platelet; hemostatic; diagnostic agent;

KW human.

XX

OS Homo sapiens.

XX

XX

EH Key Location/Qualifiers

FT Peptide 1..29

FT /note= "signal peptide"

FT Protein 30..1367

FT /note= "mature protein"

XX

PN WC9832771-A1.

XX

PD 30-JUL-1998.

XX

XX

PF 29-JAN-1998; 98WO-JP00370.

XX

PR 29-AUG-1997; 97JP-0234544.

PR 29-JAN-1997; 97JP-0015118.

XX

PA (TORA ) TORAY IND INC.

XX

PI Rainoch M, Tanaka T;

XX

DR WPI; 1998-427881/36.

DR N-PSDB; AAV33774.

XX

XX

PT Integrin-immunoglobulin chimeric protein heterodimer complexes as

PT platelet substitutes - contain the alpha and beta integrin chains

PT associated in stable state and bind to extracellular matrix in the

PT presence of plasma components

XX

XX

PS Claim 8; Pages 62-71; 87pp; Japanese.

XX

XX

CC This represents an integrin alpha-2 chain. The invention provides

CC integrin-immunoglobulin chimeric protein heterodimer complexes that

CC comprise an integrin alpha or beta chain associated with an

CC immunoglobulin light or heavy chain. These chimeric proteins form

CC heterodimer complexes, in particular with a chimeric protein containing

CC an integrin alpha chain and an immunoglobulin chain with a chimeric

CC protein containing an integrin beta chain and an immunoglobulin chain;

CC the immunoglobulin chain in each case may be a heavy chain, or one of the

CC two may be a light chain. The integrin alpha chain is preferably alpha 4

CC or alpha 2 and the integrin beta chain is preferably beta 1. Animal cells

CC transformed with vectors containing the DNA coding for the above chimeric

CC proteins can be used in the preparation of the chimeric proteins and

CC their heterodimer complexes. The heterodimer complexes, which are useful

CC for testing potential promoters and inhibitors of the binding of

CC integrins to their ligands, function as blood platelet substitutes and

CC hemostatics and as diagnostic agents.

XX

SQ Sequence 1367 AA;

Query Match 88.9%; Score 1271; DB 19; Length 1367;

Best Local Similarity 18.8%; Pred. No. 1.7e-34;

Matches 243; Conservative 7; Mismatches 14; Indels 1043; Gaps 5;

QY 1 DWL----- 3

Db 61 NLLVGSFWSGFFPENRMGDYKCPVDLSTATCEKLNLOTSTIPNVTEKTNMSLGLILT 120  
:|

QY 4 ----- 3

Db 121 RNMGTCGFLTCGPLWAQCCGNQYTTGVCSDISPDFOLSASFSPATPCPSLIDVVVVCD 180

QY 4 -----KAFYDKVAEKL----- 14  
| | : | : |

Db 181 ESNSIYPWDVAKNPFLEKFOGLDIGPTKTQVGLIQYANNPRVFNLTNTYTKTEEMIVATS 240

QY 15 ----- 14

Db 241 QTSQGGDLTNTFCAIQYARKYAYSAASSGRRSATKVMVVVTDGESHGSMKKAVIDOCN 300

QY 15 ----- 14

Db 301 HDNILRFGIAVLGYLNRNALDTKNLKEIRAIAISIPTRYFFNVSDAALKEAGTLGEQ 360

QY 15 ----- 14

Db 361 FSIEGTVQGGDNFQMEMSQVGSADYSSQNDILMLGAVGAFGWSGTIVOKTSHGLIFP 420

QY 15 KEAFD----- 19  
| : | |

Db 421 KQAFDQILODRNHSSYLGYSVAALSTGESTHFVAGAPRANTGQIVLYSVNENGNIIVIQ 480

QY 20 ----- 19

Db 481 AHRGDQIGSYFGSVLCSDVLDKDTITDVLVGPMPYMSDLKKEGRVLEFTIKKGILQGH 540

QY 20 ----- 19

Db 541 QFLEGPEGIENTRFGSAIAALSDINMGDFNDVIVGSPLENQNSGAVIYNGHQTIRTKY 600

QY 20 -----WKAFYDKVAEK- 31  
| : |

Db 601 SQKILGSDGAFRSHLOYFGRSLDGYGDLNGDSITDVSIGAGQVVLWSQSIADVAIEAS 660

QY 32 ----- 31

Db 661 FTPEKITLVNKNQAIIILKLCFSAKFRPTKONNOVAIVVINITLDADGFSRVRTSRGLFKN 720

QY 32 ----- 31

Db 721 NERCLQNMVNVNAQSCPEHIIYIQEPSDVVNSLDLRVDISLENPGTSPALEYSETAKV 780

QY 32 ----- 31

Db 781 FSIPFHKDCGEDGLCISDLVDVQIPAAQEQPFIVSNQNKRLTFSYLNKNRESAYNTG 840

QY 32 ----- 31

Db 841 IVVDFSENLEFFAGFLPVDGTEVTCQVAASQKSVACDVGYPALKREQQVFTTFINFDNLQ 900

QY 32 ----- 31

Db 901 NLNQASLSFOALSESQENKADNLVNLKIPLLYDAEIHLTRSTNINFIYESSDGNVPSI 960

QY 32 ----- 31

Db 961 VHSFEDVGKFEIFSLKVTTCGSPVPMATVLIHIPQYTKKNPLMYLTGVTQDKAGDISCN 1020

QY 32 -----LKEAFM----- 37  
| : |

Db 1021 ADINPLKIGQTSVSVFSENFHRTKELNCRNTASCNNVTCWLDVHKMGVEFVNVNTRIW 1080

```

QY 38 ----- 37
Db 1081 NGTFASSTFQVLTAAAEINTYNYPIYVIEDNTVTIPLMINKPDEKAEVPTDPEPKSC 1140
QY 38 DKTHTCPCPCAPPELLGGPSVLEFPKPKDTLMISRPETVCVVVDVSHEDPEVKFNWYVD 97
Db 1141 DKTHTCPCPCAPPELLGGPSVLEFPKPKDTLMISRPETVCVVVDVSHEDPEVKFNWYVD 1200
QY 98 GVEVHNKTKPREEQYNSTYRVVSVLTVLHQDWLNGKEYKCKVSNKALPAPIEKTSKAK 157
Db 1201 GVEVHNKTKPREEQYNSTYRVVSVLTVLHQDWLNGKEYKCKVSNKALPAPIEKTSKAK 1260
QY 158 GQPREQVYTLPPSRDELTKNQVSLTCLVKGYFSPVDIAVWESNGQPENNYKTTPPVLD 217
Db 1261 GQPREQVYTLPPSRDELTKNQVSLTCLVKGYFSPVDIAVWESNGQPENNYKTTPPVLD 1320
QY 218 DGSFFLYSKLTVDKSRWQQGNVFCSCVMHEALHNNHYTQKLSLSPGK 264
Db 1321 DGSFFLYSKLTVDKSRWQQGNVFCSCVMHEALHNNHYTQKLSLSPGK 1367

RESULT 6
AAU09817
ID AAU09817 standard; Protein; 594 AA.
XX
AC AAU09817;
XX
DT
XX
DE Murine FGFR-L extracellular domain-Fc fusion protein.
XX
KW Mouse; fibroblast growth factor receptor-like protein; FGFR-L; anorectic;
KW haemostatic; osteopathic; cytostatic; nephrotropic; antidiabetic;
KW immunomodulator; antiinflammatory; haematopoietic disorder; osteoporosis;
KW osteogenesis imperfecta; Paget's disease; periodontal disease; cancer;
KW hypercalcaemia; acute glomerulonephritis; chronic glomerulonephritis;
KW diabetes; obesity; cachexia; transgenic animal; gene therapy; ds.
XX
OS Mus musculus.
XX
FH Key
CDS 1..1785
FT /*tag= a
FT /product= "Fibroblast growth factor receptor-like protein
FT extracellular domain-Fc fusion protein"
XX
PN WO200170977-A2.
XX
PD
XX
PD 27-SEP-2001.
XX
PF 22-MAR-2001; 2001WO-US09073.
XX
PR 22-MAR-2000; 2000US-191379P.
XX
PA (AMGE-) AMGEN INC.
PA (SARI/) SARIS C M.
PA (MUSX/) MU S X.
PA (XIAM/) XIA M.
PA (BOON/) BOONE T C.
PA (COVE/) COVEY T.
XX
PI Saris CM, Mu SX, Xia M, Boone TC, Covey T;
XX
DR WPI; 2001-626128/72.
XX
DR N-PSDB; AAS14939.
XX
PT Novel nucleic acid encoding fibroblast growth factor receptor-like
PT polypeptides, useful for treating hematopoietic disorder, osteoporosis,
PT Paget's disease, glomerulonephritis, cancer, diabetes, obesity and
PT cachexia
XX
PS Example 4; Page 161-163; 163pp; English.

```

```

XX
CC The invention relates to a novel isolated fibroblast growth factor
CC receptor-like (FGFR-L) polypeptide (I). (I) and the nucleic acid (II)
CC encoding (I) are useful for treating, preventing or ameliorating
CC a medical condition including haematopoietic disorder, osteoporosis,
CC osteogenesis imperfecta, Paget's disease, periodontal disease,
CC hypercalcaemia, acute glomerulonephritis, chronic glomerulonephritis,
CC cancer, diabetes, obesity and cachexia. (I) is also useful for
CC identifying a compound which binds to FGFR-L polypeptide, by contacting
CC (I) with a compound, determining the extent of binding of the FGFR-L
CC polypeptide to the compound, and determining the activity of the
CC polypeptide when bound to the compound. (II) is useful for modulating
CC levels of a polypeptide in an animal. A transgenic animal comprising (I)
CC is useful for determining whether a compound inhibits FGFR-L polypeptide
CC activity or FGFR-L polypeptide production, by exposing the transgenic
CC animal to the compound and measuring FGFR-L polypeptide or production in
CC the animal. (II) is useful for mapping the locations of FGFR-L gene and
CC related genes on chromosomes, as hybridisation probes in diagnostic
CC assays to test for the presence of an FGFR-L nucleic acid molecule in
CC mammalian tissue or bodily fluid samples, in gene therapy, and as tools
CC for isolating corresponding FGFR-L polypeptide genes. (I) is useful as
CC immunogen, and for cloning FGFR-L polypeptide ligands using an
CC expression cloning strategy. The present sequence represents the
CC coding sequence of mouse fibroblast growth factor receptor-like protein
CC extracellular domain-Fc fusion protein used to generate anti-FGFR-L
CC antibodies as described in the invention.
XX
SQ Sequence 594 AA;
Query Match 88.7%; Score 1268; DB 22; Length 594;
Best Local Similarity 57.6%; Pred. No. 3.3e-35;
Matches 239; Conservative 9; Mismatches 15; Indels 152; Gaps 4;
QY 2 WLK----- 4
Db 180 WMKDDQTLTHLEASERKKKKWTLNKLKPEDSGYTCRVSNKAGAINATYKVDVIQRT 239
QY 5 -----AFYDKVAEKLKEAFDWLK----- 22
Db 240 SKPVLGTGTHPVNTTVDGTTSTFOCKVRSDVKPVIQWLKRVYSGEGRHNSTIDVGGQKF 299
QY 23 -----AFYDK-----VAEKLKEAFM----- 37
Db 300 VVLPTGDVMSRDPDGSYLNKLISRARQDDAGMYICLGANTMGYSRFAFLTVLPDPKPPG 359
QY 38 -----DKHTCPCPCAPPELLGGPSVLEFPKPKDTLMISRPETVCVVVDVSHEDPE 89
Db 360 PPMASSVDKTHTCPCPCAPPELLGGPSVLEFPKPKDTLMISRPETVCVVVDVSHEDPE 419
QY 90 VKFNWYVDGVEVHNKTKPREEQYNSTYRVVSVLTVLHQDWLNGKEYKCKVSNKALPAPI 149
Db 420 VKFNWYVDGVEVHNKTKPREEQYNSTYRVVSVLTVLHQDWLNGKEYKCKVSNKALPAPI 479
QY 150 EKTISKAKGQPREPQVYTLPPSRDELTKNQVSLTCLVKGYFSPVDIAVWESNGQPENNYK 209
Db 480 EKTISKAKGQPREPQVYTLPPSRDELTKNQVSLTCLVKGYFSPVDIAVWESNGQPENNYK 539
QY 210 TTPPVLDSDGSFFLYSKLTVDKSRWQQGNVFCSCVMHEALHNNHYTQKLSLSPGK 264
Db 540 TTPPVLDSDGSFFLYSKLTVDKSRWQQGNVFCSCVMHEALHNNHYTQKLSLSPGK 594

RESULT 7
AAW70798
ID AAW70798 standard; protein; 951 AA.
XX
AC AAW70798;
XX
DT 03-FEB-1999 (first entry)
XX
DE Human gp130-C-gamma-1 amino acid sequence.
XX
KW gp130; cytokine antagonist; Interleukin; gamma-Interferon;

```

KW granulocyte macrophage colony-stimulating factor; J peptide;  
XX transforming growth factor-beta.  
OS Synthetic.  
XX Homo sapiens.  
XX Key Location/Qualifiers  
FH Protein 1..619  
FT /note= "human gp130"  
FT Protein 662..651  
FT /note= "from the constant region of human IgG1"  
XX US5844099-A.  
XX 01-DEC-1998.  
XX 27-NOV-1995; 95US-0563105.  
XX 27-NOV-1995; 95US-0563105.  
PR 20-OCT-1993; 93US-0140222.  
XX (REGE-) REGENERON PHARM INC.  
XX Economides A, Stahl N, Yancopoulos GD;  
XX WPI; 1999-044669/04.  
XX Cytokine antagonists - comprising extracellular domains of  
PT specificity-determining and signal-transducing components of  
PT cytokine receptor  
XX Example 4; Fig 9A-B; 46pp; English.  
XX The present sequence represents the amino acid sequence of human  
CC gp130-C-gamma-1. The protein is used in the course of the invention.  
CC The specification describes cytokine antagonists comprising only the  
CC extracellular domain of the specificity-determining component of  
CC the cytokine receptor and the extracellular domain of a  
CC signal-transducing component of the cytokine receptor. The cytokine  
CC is an interleukin (IL-1, IL-2, IL-3, IL-4, IL-5 or IL-15),  
CC granulocyte macrophage colony-stimulating factor (GM-CSF),  
CC gamma-interferon or transforming growth factor-beta (TGF-beta). The  
CC antagonist is capable of binding the cytokine to form a nonfunctional  
CC complex. The compounds have therapeutic activity as cytokine antagonists  
CC and can also be used in assays for identifying novel agonists and  
CC antagonists of cytokines.  
XX Sequence 951 AA;  
SQ  
Query Match 88.7%; Score 1268; DB 20; Length 951;  
Best Local Similarity 38.4%; Pred. No. 9.5e-35;  
Matches 246; Conservative 8; Mismatches 10; Indels 376; Gaps 7;  
QY 1 DW-----  
DB 312 DWSEASGITYEDRPSKAPFWYKIDPSHTQGYRTVQVWVKTLPPFEANGKILDYEVTLT 371  
QY 3 DW-----  
DB 372 RWKSHLQYTVNATKLTNLTNDRLATLVNRLVGKSDAAVLTPACDFQATHPVMDLK 431  
QY 5 AF-----YDK-----VAEK-- 13  
DB 432 AFPKDNMLVWETTPRESVKYILEWCVLSDKAPCIDTQWQEDGTVHRTVLRGNLAESKC 491  
QY 14 -----LKEA-----FDW----- 20  
DB 492 YLITVTPYADGPGSPESIKAYLKQAPPKSGPTVTRTKVGKNEAVLEMDQLPDVQNGFI 551  
QY 21 -----LKAFYD----- 26  
DB 552 RNVTIFYRTIIGNETAVNVDSSHTEYTLSSLTSDTLYMVRMAAYTDEGKGDPFTFTTP 611

QY 27 ----- 26  
DB 612 KFAQGEIESGASTKGPSVFPPLAPSSKSTSGTAALGCLVKDYFPEPTVSWNSGALISGV 671  
QY 27 -----KVAEKLKEAFMDKTHTCP 44  
DB 672 HTFPAVLQSSGLYSLSSVTVTPSSSLGTQTYICNVNHPKPSNTKVDKKVEPKSCDKTHTCP 731  
QY 45 PCPAPELLGGPSVFLFPPKPKDITLMISRTPEVTCVVVDVSHEDPEVKFNWYVDGVEVHNA 104  
DB 732 PCPAPELLGGPSVFLFPPKPKDITLMISRTPEVTCVVVDVSHEDPEVKFNWYVDGVEVHNA 791  
QY 105 KTKPREEQYNSTYRVVSVLTVLHQDWLNGKEYKCKVSNKALPAPIEKTISKAKGQPREPQ 164  
DB 792 KTKPREEQYNSTYRVVSVLTVLHQDWLNGKEYKCKVSNKALPAPIEKTISKAKGQPREPQ 851  
QY 165 VYTLPPSRDELTKNQVSLTCLVKGFYPSDIAVEWESNGQPENNYKTTPPVLDSDGSFELY 224  
DB 852 VYTLPPSRDELTKNQVSLTCLVKGFYPSDIAVEWESNGQPENNYKTTPPVLDSDGSFELY 911  
QY 225 SKLTVDKSRWQQGNVFCSCVMHEALHNYHTOKSLSPGK 264  
DB 912 SKLTVDKSRWQQGNVFCSCVMHEALHNYHTOKSLSPGK 951  
RESULT 8  
AAY92186  
ID AAY92186 standard; protein; 951 AA.  
XX AC AAY92186;  
XX DT 01-AUG-2000 (first entry)  
XX DE Human gp130-C-gamma-1.  
XX KW gp130-C-gamma-1; cytokine; antagonist; CNTF; receptor; fusion protein;  
XX cytotostatic; immunomodulator; osteopathic.  
XX OS Synthetic.  
XX OS Homo sapiens.  
XX FH Key Location/Qualifiers  
FT Protein 1..619  
FT /label= gp130  
FT Peptide 620..621  
FT /note= "Ser-Gly bridge"  
FT Protein 622..951  
FT /label= IgG1\_constant\_region  
XX WO200018932-A2.  
XX PN 06-APR-2000.  
XX PD 22-SEP-1999; 99WO-US22045.  
XX PF 25-SEP-1998; 98US-0101858.  
XX PR 19-MAY-1999; 99US-0313942.  
XX PA (REGE-) REGENERON PHARM INC.  
XX PI Stahl N, Yancopoulos GD;  
XX PI WPI; 2000-293165/25.  
XX DR Isolated nucleic acid molecule for treating cytokine-related diseases  
XX or disorders encodes a fusion polypeptide capable of binding a cytokine  
PT to form a nonfunctional complex  
PT to form a nonfunctional complex  
XX Example 4; Fig 9; 152pp; English.  
XX The invention concerns production of antagonists to any cytokine that  
CC utilizes an alpha specificity determining component, which when combined  
CC with the cytokine, binds to a first beta signal transducing component to

CC form a non-functional intermediate which then binds to a second beta  
CC signal transducing component causing beta-receptor dimerization, the  
CC soluble alpha specificity determining component of the receptor,  
CC (SR-alpha) and the extracellular domain of the first beta signal  
CC transducing component of the cytokine receptor (beta-1) are combined to  
CC form heterodimers (SR-alpha:beta-1) that act as antagonist to the  
CC cytokine by binding the cytokine to form a non-functional complex. The  
CC receptor components are shared by cytokines such as the CNTF (ciliary  
CC neurotrophic factor) family of cytokines. The invention provides the  
CC basis for the development of IL-6 antagonists, as they show that if, in  
CC the presence of a ligand, a non-functional intermediate complex,  
CC consisting of the ligand, its alpha receptor and its beta-1 receptor  
CC component, can be formed, it will effectively block the action of the  
CC ligand. Effective antagonists of IL-6 or CNTF consist of heterodimers  
CC of the extracellular domains of the alpha specificity determining  
CC components of their receptors and the extracellular domain of gp130.  
CC The resultant heterodimers, function as high-affinity traps, rendering  
CC the cytokine inaccessible to form a signal transducing complex with the  
CC native membrane-bound forms of their receptor. The nucleic acids and  
CC polypeptides are useful for treating cytokine-related diseases or  
CC disorders such as osteoporosis and primary and secondary effects of  
CC cancer including multiple myeloma or cachexia.  
XX SQ Sequence 951 AA;  
Query Match 88.7%; Score 1268; DB 21; Length 951;  
Best Local Similarity 38.4%; Pred. No. 9.5e-35;  
Matches 246; Conservative 8; Mismatches 10; Indels 376; Gaps 7;  
QY 1 DW-----YDK-----VAEK-- 13  
DB 312 DWSEASGITYEDRPSKAPSFNWKIDPSHTQGYRTVOLWVKTLPPPEANGKILDYEVTLT 371  
QY 3 -----LK 4  
DB 372 RMKSHLQNTVNAKTLVNLNDRYLATLVRLNKGSDAAVLITPACDFQATHPVMDLK 431  
QY 5 AF-----YDK-----VAEK-- 13  
DB 432 AFPKDNMLVWVTPRESVKYKYLEWCVLSDKAPCIDWQEDGIVHRYTLRGNLAESKC 491  
QY 14 -----LKEA-----FDW----- 20  
DB 492 YLITVTPVYADGPGSPESIKAYLKQAPPSKGTVRTKVKGNKAEVLEMDQLPVDVONGFI 551  
QY 21 -----LKAFYD----- 26  
DB 552 RNVTIFYRTIIGNETAVNVDSSTHTYTLSSLTSDTLVYVMAAYTDEGKGQPEFTFTTP 611  
QY 27 ----- 26  
DB 612 KFAQGEIESGASTKGPSVFPPLAPSSKTSGGTAALGCLVKDYFPPEVTVSWNSGALTSV 671  
QY 27 -----KVAEKLKEAFMDKTHTCP 44  
DB 672 HTFPVAVLQSSGLYSLSSVVTPSSLSIGTYICNVNHNKPSNFKVKPEKSCDKTHTCP 731  
QY 45 PCPAPPELLGGPSVFLFPPKPKDTLMISRTPEVTCVVDVSHEDPEVKENWYDGVGVHNA 104  
DB 732 PCPAPPELLGGPSVFLFPPKPKDTLMISRTPEVTCVVDVSHEDPEVKENWYDGVGVHNA 791  
QY 105 KTKPREEQYNSTYRVSVLTVLHQDLNNGKEYKCKVSKNKPALPAIEKTIISKAKGPREPQ 164  
DB 792 KTKPREEQYNSTYRVSVLTVLHQDLNNGKEYKCKVSKNKPALPAIEKTIISKAKGPREPQ 851  
QY 165 VYTLPPSRDELTKNQVSLTCLVKGYFPPSDIAVEWESNGQPENNYKTTTPVLDSDGSFFLY 224  
DB 852 VYTLPPSRDELTKNQVSLTCLVKGYFPPSDIAVEWESNGQPENNYKTTTPVLDSDGSFFLY 911  
QY 225 SKLTVDKSRWQOGNFGVSCSVMEALHNHYVTKLSLSLSPCK 264  
DB 912 SKLTVDKSRWQOGNFGVSCSVMEALHNHYVTKLSLSLSPCK 951

RESULT 9  
AAY92187  
ID AAY92187 standard; protein; 961 AA.  
XX  
AC AAY92187;  
XX  
DT 01-AUG-2000 (first entry)  
XX  
DE Human gp130-J-C-gamma-1.  
XX  
KW gp130-J-C-gamma-1; cytokine; antagonist; CNTF; receptor; fusion protein;  
KW cytostatic; immunomodulator; osteopathic.  
XX  
OS Synthetic.  
OS Homo sapiens.  
XX  
FH Key Location/Qualifiers  
FT Protein 1..619  
FT /label= gp130  
FT Peptide 620..621  
FT /note= "Ser-Gly bridge"  
FT Peptide 622..631  
FT /note= "J-peptide"  
FT Protein 632..961  
FT /label= IgG1\_constant\_region  
XX  
PN WO200018932-A2.  
XX  
PD 06-APR-2000.  
XX  
PF 22-SEP-1999; 99WO-US22045.  
XX  
PR 25-SEP-1998; 98US-0101858.  
PR 19-MAY-1999; 99US-0313942.  
XX  
PA (REGE-) REGENERON PHARM INC.  
XX  
PI Stahl N, Yancopoulos GD;  
XX  
DR WPI; 2000-293165/25.  
XX  
PT Isolated nucleic acid molecule for treating cytokine-related diseases  
PT or disorders encodes a fusion polypeptide capable of binding a cytokine  
PT to form a nonfunctional complex  
XX  
PS Example 4; Page -: 152pp; English.  
XX  
CC The invention concerns production of antagonists to any cytokine that  
CC utilizes an alpha specificity determining component, which when combined  
CC with the cytokine, binds to a first beta signal transducing component to  
CC form a non-functional intermediate which then binds to a second beta  
CC signal transducing component causing beta-receptor dimerization, the  
CC soluble alpha specificity determining component of the receptor  
CC (SR-alpha) and the extracellular domain of the first beta signal  
CC transducing component of the cytokine receptor (beta-1) are combined to  
CC form heterodimers (SR-alpha:beta-1) that act as antagonist to the  
CC cytokine by binding the cytokine to form a non-functional complex. The  
CC receptor components are shared by cytokines such as the CNTF (ciliary  
CC neurotrophic factor) family of cytokines. The invention provides the  
CC basis for the development of IL-6 antagonists, as they show that if, in  
CC the presence of a ligand, a non-functional intermediate complex,  
CC consisting of the ligand, its alpha receptor and its beta-1 receptor  
CC component, can be formed, it will effectively block the action of the  
CC ligand. Effective antagonists of IL-6 or CNTF consist of heterodimers  
CC of the extracellular domains of the alpha specificity determining  
CC components of their receptors and the extracellular domain of gp130.  
CC The resultant heterodimers, function as high-affinity traps, rendering  
CC the cytokine inaccessible to form a signal transducing complex with the  
CC native membrane-bound forms of their receptor. The nucleic acids and  
CC polypeptides are useful for treating cytokine-related diseases or  
CC disorders such as osteoporosis and primary and secondary effects of  
CC cancer including multiple myeloma or cachexia.  
CC





Db 806 TPVYADGPGSPESIKAYLKQAPPSKGTPTVTKKVGKNEAVLEWDQLPVDVQNGFIRNYTI 865  
QY 38 -----  
Db 866 FYRTIIGNEAVNVDSSTHTYTLSSLTSDLYMVRMAAYTDEGKGDPETFTTPKFAQG 925  
QY 38 -----DKTHTCCPCPAPELLGGPSVFLFPKPKDTLMISRTPEVTCVVVDVSHEDPEVK 91  
Db 926 EIESGGDKTHTCCPCPAPELLGGPSVFLFPKPKDTLMISRTPEVTCVVVDVSHEDPEVK 985  
QY 92 FNMVYDGVVEVHNAKTPREEQYNSTYRVVSVLTVLHQDWLNGKEYKCKVSNKALPAPIEK 151  
Db 986 FNMVYDGVVEVHNAKTPREEQYNSTYRVVSVLTVLHQDWLNGKEYKCKVSNKALPAPIEK 1045  
QY 152 TISKAKGQPREPQVYTLPPSRDELTKNQVSLTCLVKGFYPSDIAVEHESNGQPENNYKTT 211  
Db 1046 TISKAKGQPREPQVYTLPPSRDELTKNQVSLTCLVKGFYPSDIAVEHESNGQPENNYKTT 1105  
QY 212 PPVLDSGSGFFLYSKLTVDKSRWQOGNVFSCVMHEALHNHYTKLSLSPGK 264  
Db 1106 PPVLDSGSGFFLYSKLTVDKSRWQOGNVFSCVMHEALHNHYTKLSLSPGK 1158

RESULT 11  
AAY92204  
ID AAY92204 standard; Protein; 1168 AA.  
XX  
AC AAY92204;  
XX  
DT 01-AUG-2000 (first entry)  
XX  
DE Fusion polypeptide 412, IL-6 trap.  
XX  
KW IL-6 trap; cytokine; antagonist; CNTF; receptor; fusion protein;  
KW cytostatic; immunomodulator; osteopathic.  
XX  
OS Synthetic.  
OS Homo sapiens.  
PN WO200018932-A2.  
PD 06-APR-2000.  
XX  
PF 22-SEP-1999; 99WO-US22045.  
XX  
PR 25-SEP-1998; 98US-0101858.  
PR 19-MAY-1999; 99US-0313942.  
XX  
PA (REG-) REGENERON PHARM INC.  
XX  
PI Stahl N, Yancopoulos GD;  
XX  
PI WPI: 2000-293165/25.  
DR N-PSDB; AAA09046.  
XX  
PT Isolated nucleic acid molecule for treating cytokine-related diseases  
PT or disorders encodes a fusion polypeptide capable of binding a cytokine  
PT to form a nonfunctional complex  
XX  
PS Example 6; Fig 24A-F; 152pp; English.  
XX  
CC This sequence shows fusion polypeptide 412, which is capable of  
CC binding cytokine IL-6 to form a non-functional complex.  
CC The invention concerns production of antagonists to any cytokine that  
CC utilizes an alpha specificity determining component, which when combined  
CC with the cytokine, binds to a first beta signal transducing component to  
CC form a non-functional intermediate which then binds to a second beta  
CC signal transducing component causing beta-receptor dimerization, the  
CC soluble alpha specificity determining component of the receptor  
CC (sr-alpha) and the extracellular domain of the first beta signal  
CC transducing component of the cytokine receptor (beta-1) are combined to  
CC form heterodimers (sr-alpha:beta-1) that act as antagonist to the  
CC cytokine by binding the cytokine to form a non-functional complex. The

CC receptor components are shared by cytokines such as the CNTF (ciliary  
CC neurotrophic factor) family of cytokines. The invention provides the  
CC basis for the development of IL-6 antagonists, as they show that if, in  
CC the presence of a ligand, a non-functional intermediate complex,  
CC consisting of the ligand, its alpha receptor and its beta-1 receptor  
CC component, can be formed, it will effectively block the action of the  
CC ligand. Effective antagonists of IL-6 or CNTF consist of heterodimers  
CC of the extracellular domains of the alpha specificity determining  
CC components of their receptors and the extracellular domain of gp130.  
CC The resultant heterodimers, function as high-affinity traps, rendering  
CC the cytokine inaccessible to form a signal transducing complex with the  
CC native membrane-bound forms of their receptor. The nucleic acids and  
CC polypeptides are useful for treating cytokine-related diseases or  
CC disorders such as osteoporosis and primary and secondary effects of  
CC cancer including multiple myeloma or cachexia.  
XX  
SQ Sequence 1168 AA;  
Query Match 88.7%; Score 1268; DB 21; Length 1168;  
Best Local Similarity 37.1%; Pred. No 1.5e-34;  
Matches 242; Conservative 12; Mismatches 9; Indels 390; Gaps 6;  
QY 2 WLKA-----FYDKV-----AEKL----- 14  
Db 516 WVEAENALGKVTSDHINFDPYKVPKPNPHNLSVINSEELSLKLTWTNPSIKSVILK 575  
QY 15 -----KEAFDW----- 20  
Db 576 YNIOYRTKDASTWSQIPPEDTASTRSFTVODLKPTEYVFRICMKEDGKGYWSDSEE 635  
QY 21 ----- 20  
Db 636 ASGITVEDRPSKAPSWYKIDPSHTOGYRTVQLVWKTLPPEFANGKILDYEVTLRWKSH 695  
QY 21 -----LKAF--- 24  
Db 696 LQNYTVNATKLTNLTNDRLATLTVRNLVGKSDAAVLTIPACDFQATHPVMDLKAFPKD 755  
QY 25 ----- 24  
Db 756 NMLWVETTPRESVKKYILEWCVLSDKAPCITDMQEDGTVHRTYLRGNLAESKYLITV 815  
QY 25 -----YDKVAEKLKEAFM----- 37  
Db 816 TPVYADGPGSPESIKAYLKQAPPSKGTPTVTKKVGKNEAVLEWDQLPVDVQNGFIRNYTI 875  
QY 38 ----- 37  
Db 876 FYRTIIGNEAVNVDSSTHTYTLSSLTSDLYMVRMAAYTDEGKGDPETFTTPKFAQG 935  
QY 38 -----DKTHTCCPCPAPELLGGPSVFLFPKPKDTLMISRTPEVTCVVVDVSHEDPEVK 91  
Db 936 EIESGGDKTHTCCPCPAPELLGGPSVFLFPKPKDTLMISRTPEVTCVVVDVSHEDPEVK 995  
QY 92 FNMVYDGVVEVHNAKTPREEQYNSTYRVVSVLTVLHQDWLNGKEYKCKVSNKALPAPIEK 151  
Db 996 FNMVYDGVVEVHNAKTPREEQYNSTYRVVSVLTVLHQDWLNGKEYKCKVSNKALPAPIEK 1055  
QY 152 TISKAKGQPREPQVYTLPPSRDELTKNQVSLTCLVKGFYPSDIAVEHESNGQPENNYKTT 211  
Db 1056 TISKAKGQPREPQVYTLPPSRDELTKNQVSLTCLVKGFYPSDIAVEHESNGQPENNYKTT 1115  
QY 212 PPVLDSGSGFFLYSKLTVDKSRWQOGNVFSCVMHEALHNHYTKLSLSPGK 264  
Db 1116 PPVLDSGSGFFLYSKLTVDKSRWQOGNVFSCVMHEALHNHYTKLSLSPGK 1168  
RESULT 12  
AAG63640  
ID AAG63640 standard; Protein; 475 AA.  
XX  
AC AAG63640;  
XX



CC disorders such as osteoporosis and primary and secondary effects of  
XX cancer including multiple myeloma or cachexia.  
SQ Sequence 690 AA;

Query Match 88.5%; Score 1265; DB 21; Length 690;  
Best Local Similarity 42.7%; Pred. No. 5.9e-35;  
Matches 238; Conservative 9; Mismatches 17; Indels 294; Gaps 3;

Qy 1 DW-----LKAFLDKVAEKLK----- 16  
Db 133 EWGPRSTSLTKAVLLVRKQNSPAEDFQPCQYSQSFSCOLAYPEGDSFFIYISM 192  
Qy 17 -----AFDWLKAFLK----- 25  
Db 193 CVASSVGSKFTQFGGILQPPDPPANIITVAVARNRMLSVTWQDPHNSNSFYRLR 252  
Qy 26 ----- 25  
Db 253 FELRYAERSKFTFTTMMVKDLQHHCVIHDWAGSLRHVVQLRAQERFGQSEWSPEANG 312  
Qy 26 ----- 25  
Db 313 TPWTERSPPAENESTVPMQALTTNNKDDNIFRDSANATSLPVODAGASTGSPVFPLA 372  
Qy 26 ----- 25  
Db 373 PSSKSTSGTAALGCLVKDYFPEPVTVSWNSGALTSVHTFPAVLQSSGLYSLSSVVTVP 432  
Qy 26 -----DKVAEKLKAFMDKHTCCPCPAPELLGGPSVFLPPPKD 66  
Db 433 SSSLTQTYICNVNHPKNTKVDKKEPKSCDKHTCCPCPAPELLGGPSVFLPPPKD 492  
Qy 67 TLMISRTPEVTCVVVDVSHEDPEVKFNWYVDGVEVHNAKTKPREEQYNSTYRVVSVLTVL 126  
Db 493 TLMISRTPEVTCVVVDVSHEDPEVKFNWYVDGVEVHNAKTKPREEQYNSTYRVVSVLTVL 552  
Qy 127 HDWLNKREYCKVSNKALPAPIEKTISKAKGQPREPQVYTLPPSRDELTKNQVSLTCLV 186  
Db 553 HDWLNKREYCKVSNKALPAPIEKTISKAKGQPREPQVYTLPPSRDELTKNQVSLTCLV 612  
Qy 187 KGFYPSDIAVWESNGQPENNYKTTPPVLDSDGSFFLYSKLTVDKSRWQQGNVSCSVMH 246  
Db 613 KGFYPSDIAVWESNGQPENNYKTTPPVLDSDGSFFLYSKLTVDKSRWQQGNVSCSVMH 672  
Qy 247 EALHNHYTQKSLSLSPGK 264  
Db 673 EALHNHYTQKSLSLSPGK 690

RESULT 14  
AAG64473  
ID AAG64473 standard; Protein: 473 AA.  
XX  
AC AAG64473;  
XX  
DT 25-SEP-2001 (first entry)  
XX  
DE Human type antihuman IgE antibody H chain 3.  
XX  
KW Human; antihuman IgE antibody; immunoglobulin; treating;  
KW allergic disease.  
XX  
OS Homo sapiens.  
XX  
FH Key Location/Qualifiers  
FT Misc-difference 22  
FT /note= "Encoded by CAG"  
FT Misc-difference 129..141  
FT /note= "Encoded by ccgtggggcc agggaaacacc ggtgccttt  
FT gactacgtc"  
XX  
PN W0200151507-AL.

XX 19-JUL-2001.  
PD  
XX 15-JAN-2001; 2001WO-JP00181.  
PF  
XX 14-JAN-2000; 2000JP-0007061.  
PR  
XX (SNOW ) SNOW BRAND MILK PROD CO LTD.  
PA  
XX Washida N, Takahashi K, Satake T, Fujise N, Tanaka H, Kuriyama M;  
PI WPI; 2001-442132/47.  
XX N-PSDB; AAH47901.  
DR  
XX New peptide used for screening human anti-human immunoglobulin E  
PT monoclonal antibody useful in medical compositions for treating  
PT allergies -  
XX  
XX Example 6; Page 59-60; 70pp; Japanese.  
PS  
XX The present sequence is that of a human type antihuman IgE antibody H  
CC chain. The invention relates to a peptide useful in a method for  
CC screening for human type antihuman IgE monoclonal antibodies  
CC (AAH47897-AAH47904 encoding AAG64469-AAG64476) useful for preventing  
CC and/or treating allergic disease.  
XX  
SQ Sequence 473 AA;

Query Match 88.4%; Score 1264; DB 22; Length 473;  
Best Local Similarity 51.1%; Pred. No. 2.8e-35;  
Matches 241; Conservative 8; Mismatches 15; Indels 208; Gaps 4;

Qy 1 DWLKAFLDKVA-----EKL-----KEAFDWLK----- 22  
Db 2 DWTWRFLEFVAAATGVSQBELVQSGAEVKOTGSSRVSCRASGGTFSRYALNWVRVPG 61  
Qy 23 -----A 23  
Db 62 QGLEWMAGIIPFGPPKYAQFGVRVSLTADRSTNTAYMENARLRSDDTAVVYCARPTP 121  
Qy 24 FYD----- 26  
Db 122 LYDTETGPDYWGOGTPVAVSSASTGKSPVFLAPSSKSTSGGTAALGCLVKDYFPEPT 181  
Qy 27 -----KVAEKL 32  
Db 182 VSWNSGALTSVHTFPAVLQSSGLYSLSSVTVVPSSSLGTQTYICNVNHPKNTKVDK 241  
Qy 33 KEAFMDKHTCCPCPAPELLGGPSVFLPPPKPKDTLMISRTPEVTCVVVDVSHEDPEVK 92  
Db 242 EPKSCDKHTCCPCPAPELLGGPSVFLPPPKPKDTLMISRTPEVTCVVVDVSHEDPEVK 301  
Qy 93 NWYVDGVEVHNAKTKPREEQYNSTYRVVSVLTVLHDWLNKREYCKVSNKALPAPIEKT 152  
Db 302 NWYVDGVEVHNAKTKPREEQYNSTYRVVSVLTVLHDWLNKREYCKVSNKALPAPIEKT 361  
Qy 153 ISKAKGQPREPQVYTLPPSRDELTKNQVSLTCLVKGFYPSDIAVWESNGQPENNYKTTP 212  
Db 362 ISKAKGQPREPQVYTLPPSRDELTKNQVSLTCLVKGFYPSDIAVWESNGQPENNYKTTP 421  
Qy 213 PVLDSDGSFFLYSKLTVDKSRWQQGNVSCSVMHEALHNHYTQKSLSLSPGK 264  
Db 422 PVLDSDGSFFLYSKLTVDKSRWQQGNVSCSVMHEALHNHYTQKSLSLSPGK 473

RESULT 15  
AAG64475  
ID AAG64475 standard; Protein: 473 AA.  
XX  
AC AAG64475;  
XX  
DT 25-SEP-2001 (first entry)  
XX





GenCore version 5.1.4.p5.4578  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: April 21, 2003, 10:42:26 ; Search time 14.7529 seconds  
(without alignments)  
526.515 Million cell updates/sec

Title: 7LINK7LINK2

Perfect score: 1430

Sequence: 1 DMLKAFYDKVAEKLEAFDW.....MHEALHNHYTKLSLSLSPCK 264

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.0

Searched: 262574 seqs, 29422922 residues

Total number of hits satisfying chosen parameters: 262574

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued\_Patents\_AA.\*

- 1: /cgn2\_6/ptodata/1/iaa/5A\_COMB.pep.\*
- 2: /cgn2\_6/ptodata/1/iaa/5B\_COMB.pep.\*
- 3: /cgn2\_6/ptodata/1/iaa/6A\_COMB.pep.\*
- 4: /cgn2\_6/ptodata/1/iaa/6B\_COMB.pep.\*
- 5: /cgn2\_6/ptodata/1/iaa/PTUS\_COMB.pep.\*
- 6: /cgn2\_6/ptodata/1/iaa/backfiles1.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1261	88.2	476	3	US-08-487-550-12
2	1260	88.1	449	1	US-08-458-516-13
3	1259	88.0	664	3	US-08-957-063-16
4	1259	88.0	664	3	US-08-957-063-18
5	1259	88.0	664	4	US-09-487-685-16
6	1259	88.0	664	4	US-09-487-685-18
7	1259	88.0	664	4	US-08-802-805D-16
8	1259	88.0	664	4	US-08-802-805D-18
9	1258	88.0	476	2	US-08-378-939-10
10	1257	87.9	437	5	PCT-US96-10043-11
11	1257	87.9	454	2	US-07-934-373C-22
12	1257	87.9	454	3	US-08-437-642B-22
13	1257	87.9	454	4	US-08-146-206C-22
14	1257	87.9	454	5	PCT-US93-07832-22
15	1256	87.8	347	1	US-07-940-861-43
16	1256	87.8	347	1	US-08-459-512-43
17	1256	87.8	347	2	US-08-459-657-43
18	1256	87.8	347	2	US-08-460-132-43
19	1256	87.8	347	4	US-08-466-465-8
20	1256	87.8	347	5	PCT-US92-02050-43
21	1255	87.8	452	3	US-09-027-449-71
22	1255	87.8	452	4	US-09-026-985-71
23	1255	87.8	452	4	US-09-121-952A-71
24	1255	87.8	452	4	US-09-234-340A-71
25	1254	87.7	459	1	US-08-157-101A-7
26	1252	87.6	424	5	PCT-US95-03865-12
27	1252	87.6	424	5	PCT-US95-03866-14

28	1252	87.6	482	4	US-09-189-129-2	Sequence 2, Appli
29	1251	87.5	453	3	US-08-466-151-8	Sequence 8, Appli
30	1251	87.5	453	4	US-08-466-163B-8	Sequence 8, Appli
31	1251	87.5	467	4	US-09-301-593-18	Sequence 18, Appli
32	1251	87.5	472	4	US-09-049-672A-8	Sequence 8, Appli
33	1251	87.5	472	4	US-09-301-593-30	Sequence 30, Appli
34	1251	87.5	911	2	US-08-484-438-10	Sequence 10, Appli
35	1250	87.4	446	3	US-08-397-411-7	Sequence 7, Appli
36	1250	87.4	450	4	US-09-532-856-5	Sequence 5, Appli
37	1250	87.4	476	3	US-08-487-550-4	Sequence 4, Appli
38	1250	87.4	977	4	US-09-590-656-1	Sequence 1, Appli
39	1249	87.3	478	3	US-08-487-550-8	Sequence 8, Appli
40	1248	87.3	388	4	US-09-131-247-16	Sequence 16, Appli
41	1248	87.3	442	5	PCT-US96-10043-9	Sequence 9, Appli
42	1248	87.3	451	2	US-08-887-352B-14	Sequence 14, Appli
43	1248	87.3	451	2	US-08-887-352B-16	Sequence 16, Appli
44	1248	87.3	451	3	US-08-466-151-65	Sequence 65, Appli
45	1248	87.3	451	4	US-09-109-207C-14	Sequence 14, Appli

ALIGNMENTS

RESULT 1

US-08-487-550-12

; Sequence 12, Application US/08487550

; Patent No. 6113898

; GENERAL INFORMATION:

; APPLICANT: Anderson, Darrell R.

; TITLE OF INVENTION: "MONKEY MONOCLONAL ANTIBODIES SPECIFIC

; TITLE OF INVENTION: TO HUMAN B7.1 AND/OR B7.2 PRIMATIZED FORMS THEREOF,

; TITLE OF INVENTION: PHARMACEUTICAL COMPOSITIONS CONTAINING, AND USE THEREOF AS

; TITLE OF INVENTION: IMMUNOSUPPRESSANTS"

; NUMBER OF SEQUENCES: 12

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: BURNS, DOANE, SWECKER & MATHIS

; STREET: 699 Prince Street

; CITY: Alexandria

; STATE: VA

; COUNTRY: USA

; ZIP: 22314

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: PatentIn Release #1.0, Version #1.30

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/08/487,550

; FILING DATE: 07-JUN-1995

; CLASSIFICATION: 435

; ATTORNEY/AGENT INFORMATION:

; NAME: Teskin, Robin L.

; REGISTRATION NUMBER: 35,030

; REFERENCE/DOCKET NUMBER: 012712-131

; TELEPHONE: 703-836-6620

; TELEFAX: 703-836-2021

; INFORMATION FOR SEQ ID NO: 12:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 476 amino acids

; TYPE: amino acid

; TOPOLOGY: linear

; MOLECULE TYPE: protein

; US-08-487-550-12

Query Match 88.2%; Score 1261; DB 3; Length 476;

Best Local Similarity 53.0%; Pred. No. 1.3e-41;

Matches 236; Conservative 7; Mismatches 13; Indels 189; Gaps 2;

QY 9 KVAEKLK-----EAPDWLKAFYD-----26

| : | |

Db 32 KPSETLSLCAVSGSGISGGYWGWIWPPKGLGIEWGFSYSSGNTYNPGLKSLKSVTIS 91

QY 27 ----- 26  
Db 92 TDTSKNQFSLKLSMTAADTAVYICVRDRLESVVGVMYNNWEDVWGPGLVTVSSASTKG 151  
QY 27 ----- 26  
Db 152 PSVFPLAPSSKSTSGTAAALGCLVKDKYFPEPVTVSNWNSGALTSVHTFPAVLQSSGLYSL 211  
QY 27 ----- KVAEKLKEAFMDKTHTCPPCPAPELLGGPSVFL 59  
Db 212 SSVVTVPSSSLGTQTYICNVNHNKPSNTKDKKAEKPCDKTHTCPPCPAPELLGGPSVFL 271  
QY 60 FPPKPKDTLMISRTPEVTCVVDVSHEDPEVKFNWYVDGVEVHNKTKPREQYNSTYRV 119  
Db 272 FPPKPKDTLMISRTPEVTCVVDVSHEDPEVKFNWYVDGVEVHNKTKPREQYNSTYRV 331  
QY 120 VSVLTVLHODWLNKGEYKCKVSNKALPAPIEKTISKAKGPREPQVYTLPPSRDELTKNQ 179  
Db 332 VSVLTVLHODWLNKGEYKCKVSNKALPAPIEKTISKAKGPREPQVYTLPPSRDELTKNQ 391  
QY 180 VSLTCLVKGFYPSDIAVEWESNGQPENNYKTTPPVLDSDGSFFLYSKLTVDKSRWQQGNV 239  
Db 392 VSLTCLVKGFYPSDIAVEWESNGQPENNYKTTPPVLDSDGSFFLYSKLTVDKSRWQQGNV 451  
QY 240 FSCVMHEALHNHYTKQSLSPGK 264  
Db 452 FSCVMHEALHNHYTKQSLSPGK 476

RESULT 2

US-08-458-516-13  
; Sequence 13, Application US/08458516  
; Patent No. 5777085  
; GENERAL INFORMATION:  
; APPLICANT: Co, Man Sung  
; APPLICANT: TSO, J. Yun  
; TITLE OF INVENTION: Humanized Antibodies Reactive with  
; TITLE OF INVENTION: GPIIB/IIIA  
; NUMBER OF SEQUENCES: 23  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: William M. Smith  
; STREET: One Market Plaza, Steuart Tower, Suite 2000  
; CITY: San Francisco  
; STATE: California  
; COUNTRY: USA  
; ZIP: 94105  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/458,516  
; FILING DATE:  
; CLASSIFICATION: 424  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 08/059,159  
; FILING DATE: 03-MAY-1993  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Smith, William M.  
; REGISTRATION NUMBER: 30,223  
; REFERENCE/DOCKET NUMBER: 11823-37-3  
; TELEPHONE: 415-326-2400  
; TELEFAX: 415-326-2422  
; INFORMATION FOR SEQ ID NO: 13:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 449 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
US-08-458-516-13

Query Match 88.1%; Score 1260; DB 1; Length 449;  
Best Local Similarity 55.9%; Pred. No. 1.3e-41;  
Matches 236; Conservative 10; Mismatches 14; Indels 162; Gaps 2;  
QY 5 AFYDKVAKLKEA-----FDWLKAFY----- 25  
Db 28 AFTNVLIEWVRQAPGQLEWIGVIYPGSGGTNYNEKFKGRVTLTVDESTNTAYMELSSLR 87  
QY 26 ----- 25  
Db 88 SEDTAVYFCARRDGNWGNFAYWQGTLTVTVSSASTKGPSVFFPLAPSSKSTSGTAAALGCL 147  
QY 26 ----- 25  
Db 148 VKDYFPEPVTVSNWNSGALTSVHTFPAVLQSSGLYSLSSVTVVPSSSLGTQTYICNVNHNK 207  
QY 26 ---DKVAEKLKEAFMDKTHTCPPCPAPELLGGPSVFLFPPKPKDTLMISRTPEVTCVVVD 82  
Db 208 PSNTKVDKKVEPKSCDKTHTCPPCPAPELLGGPSVFLFPPKPKDTLMISRTPEVTCVVVD 267  
QY 83 VSHEDPEVKFNWYVDGVEVHNKTKPREQYNSTYRVVSVLTVLHODWLNKGEYKCKVSN 142  
Db 268 VSHEDPEVKFNWYVDGVEVHNKTKPREQYNSTYRVVSVLTVLHODWLNKGEYKCKVSN 327  
QY 143 KALPAPIEKTISKAKGPREPQVYTLPPSRDELTKNQVSLTCLVKGFYPSDIAVEWESNG 202  
Db 328 KALPAPIEKTISKAKGPREPQVYTLPPSRDELTKNQVSLTCLVKGFYPSDIAVEWESNG 387  
QY 203 OPENNYKTTPPVLDSDGSFFLYSKLTVDKSRWQQGNVSCSVNHEALHNHYTKQSLSP 262  
Db 388 OPENNYKTTPPVLDSDGSFFLYSKLTVDKSRWQQGNVSCSVNHEALHNHYTKQSLSP 447  
QY 263 GK 264  
Db 448 GK 449  
RESULT 3  
US-08-957-063-16  
; Sequence 16, Application US/08957063  
; Patent No. 6025157  
; GENERAL INFORMATION:  
; APPLICANT: Robert D. Klein, Arnon Rosenthal, Mary A. Hynes  
; TITLE OF INVENTION: Neurturin Receptor  
; NUMBER OF SEQUENCES: 19  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Genentech, Inc.  
; STREET: 1 DNA Way  
; CITY: South San Francisco  
; STATE: California  
; COUNTRY: USA  
; ZIP: 94080  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: 3.5 inch, 1.44 Mb floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: WinPatIn (Genentech)  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/957,063  
; FILING DATE: 24-Oct-1997  
; CLASSIFICATION: 800  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 08/871  
; FILING DATE: 9-Jun-1997  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 913  
; FILING DATE: 18-Feb-1997  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Torchia, PhD., Timothy E.  
; REGISTRATION NUMBER: 36,700  
; REFERENCE/DOCKET NUMBER: P1086P2  
; TELECOMMUNICATION INFORMATION:



TELEPHONE: 650/225-8674  
TELEFAX: 650/952-9881  
INFORMATION FOR SEQ ID NO: 16:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 664 amino acids  
TYPE: Amino Acid  
TOPOLOGY: Linear  
US-08-957-063-16

Query Match 88.0%; Score 1259; DB 3; Length 664;  
Best Local Similarity 51.4%; Pred. No. 3.6e-41;  
Matches 237; Conservative 11; Mismatches 14; Indels 199; Gaps 4;

```
QY 3 LKAFYDKV-----AEKLKE-----16
||:|:|:|
DB 204 LROFFDRVPSEYTYRMLFCSCQOQACAEARRRQTILPSCSYEDKEKPNCLDLRGVCRDHL 263
||:|:|:|
QY 17 -----AFDWLKAFFD-----26
DB 264 CRSRLADFHANCRASYQTVTSCPADNYQACLSGYAGMIGDMTPNYVDSSPTGIVVSPWC 323
||:|:|:|
QY 27 -----KVAEKLKEAF-----36
DB 324 SCRGSGNMEEECEKFLDRFTENPCLRNAIQAFNGTGVNVPKGPSQATQAPRVEKTPS 383
||:|:|:|
QY 37 -----MDKTHTC 43
DB 384 LPDDLSDSTSLGTSVITTCSTVOBQGLKANNKSKLSMCFTELTTNIIPGRDPVDKTHTC 443
||:|:|:|
QY 44 PPCAPPELLGSGPSVFLPPPKPKDTLMISRTPEVTCVVVDVSHEDPEVKFNWYVDGVEVHN 103
||:|:|:|
DB 444 PPCAPPELLGSGPSVFLPPPKPKDTLMISRTPEVTCVVVDVSHEDPEVKFNWYVDGVEVHN 503
||:|:|:|
QY 104 AKTKPREEQNSTYRVVSVTLVHLDWLNKGEYKCKVSNKALPAPIEKTISKAKGQPREP 163
||:|:|:|
DB 504 AKTKPREEQNSTYRVVSVTLVHLDWLNKGEYKCKVSNKALPAPIEKTISKAKGQPREP 563
||:|:|:|
QY 164 QVYTLPPSRDELTKNOVSLTCLVKGFYPSDIAVWESNGQPENNYKTPPPVLDSDGSFFL 223
||:|:|:|
DB 564 QVYTLPPSRDEMTKNOVSLTCLVKGFYPSDIAVWESNGQPENNYKTPPPVLDSDGSFFL 623
||:|:|:|
QY 224 YSKLTVDKSRWQOGNVEFSCVMHEALHNHYTQKSLSLSPGK 264
DB 624 YSKLTVDKSRWQOGNVEFSCVMHEALHNHYTQKSLSLSPGK 664
||:|:|:|
```

RESULT 4  
US-08-957-063-18

Sequence 18, Application US/08957063  
Patent No. 6025157  
GENERAL INFORMATION:  
APPLICANT: Robert D. Klein, Arnon Rosenthal, Mary A. Hynes  
TITLE OF INVENTION: Neurturin Receptor  
NUMBER OF SEQUENCES: 19  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Genentech, Inc.  
STREET: 1 DNA Way  
CITY: South San Francisco  
STATE: California  
COUNTRY: USA  
ZIP: 94080  
COMPUTER READABLE FORM:  
MEDIUM TYPE: 3.5 inch, 1.44 Mb floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: WinPatIn (Genentech)  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/957,063  
FILING DATE: 24-Oct-1997  
CLASSIFICATION: 800  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/871  
FILING DATE: 9-Jun-1997

PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 913  
FILING DATE: 18-Feb-1997  
ATTORNEY/AGENT INFORMATION:  
NAME: Torchia, PhD., Timothy E.  
REGISTRATION NUMBER: 36,700  
REFERENCE/DOCKET NUMBER: P1086P2  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 650/225-8674  
TELEFAX: 650/952-9881  
INFORMATION FOR SEQ ID NO: 18:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 664 amino acids  
TYPE: Amino Acid  
TOPOLOGY: Linear  
US-08-957-063-18

Query Match 88.0%; Score 1259; DB 3; Length 664;  
Best Local Similarity 51.4%; Pred. No. 3.6e-41;  
Matches 237; Conservative 11; Mismatches 14; Indels 199; Gaps 4;

```
QY 3 LKAFYDKV-----AEKLKE-----16
||:|:|:|
DB 204 LROFFDRVPSEYTYRMLFCSCQOQACAEARRRQTILPSCSYEDKEKPNCLDLRSLCRTDHL 263
||:|:|:|
QY 17 -----AFDWLKAFFD-----26
DB 264 CRSRLADFHANCRASYRTITSCPADNYQACLSGYAGMIGDMTPNYVDSSPTGIVVSPWC 323
||:|:|:|
QY 27 -----KVAEKLKEAF-----36
DB 324 NCRSGNMEEECEKFLDRFTENPCLRNAIQAFNGTGVNVPKGPSLPAQAPRVEKTPS 383
||:|:|:|
QY 37 -----MDKTHTC 43
DB 384 LPDDLSDSTSLGTSVITTCSTIOBQGLKANNKSKLSMCFTELTTNISPSGRDPVDKTHTC 443
||:|:|:|
QY 44 PPCAPPELLGSGPSVFLPPPKPKDTLMISRTPEVTCVVVDVSHEDPEVKFNWYVDGVEVHN 103
||:|:|:|
DB 444 PPCAPPELLGSGPSVFLPPPKPKDTLMISRTPEVTCVVVDVSHEDPEVKFNWYVDGVEVHN 503
||:|:|:|
QY 104 AKTKPREEQNSTYRVVSVTLVHLDWLNKGEYKCKVSNKALPAPIEKTISKAKGQPREP 163
||:|:|:|
DB 504 AKTKPREEQNSTYRVVSVTLVHLDWLNKGEYKCKVSNKALPAPIEKTISKAKGQPREP 563
||:|:|:|
QY 164 QVYTLPPSRDELTKNOVSLTCLVKGFYPSDIAVWESNGQPENNYKTPPPVLDSDGSFFL 223
||:|:|:|
DB 564 QVYTLPPSRDEMTKNOVSLTCLVKGFYPSDIAVWESNGQPENNYKTPPPVLDSDGSFFL 623
||:|:|:|
QY 224 YSKLTVDKSRWQOGNVEFSCVMHEALHNHYTQKSLSLSPGK 264
DB 624 YSKLTVDKSRWQOGNVEFSCVMHEALHNHYTQKSLSLSPGK 664
||:|:|:|
```

RESULT 5

US-09-487-685-16  
Sequence 16, Application US/09487685  
Patent No. 6342348  
GENERAL INFORMATION:  
APPLICANT: Robert D. Klein, Arnon Rosenthal, Mary A. Hynes  
TITLE OF INVENTION: Neurturin Receptor  
NUMBER OF SEQUENCES: 19  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Genentech, Inc.  
STREET: 1 DNA Way  
CITY: South San Francisco  
STATE: California  
COUNTRY: USA  
ZIP: 94080  
COMPUTER READABLE FORM:  
MEDIUM TYPE: 3.5 inch, 1.44 Mb floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS

```
; SOFTWARE: WinPatin (Genentech)
; CURRENT APPLICATION DATA:
;   APPLICATION NUMBER: US/09/487,685
;   FILING DATE: 19-Jan-2000
;   CLASSIFICATION: <unknown>
; PRIOR APPLICATION DATA:
;   APPLICATION NUMBER: 08/957,063
;   FILING DATE: <unknown>
;   APPLICATION NUMBER: 913
;   FILING DATE: 18-Feb-1997
; ATTORNEY/AGENT INFORMATION:
;   NAME: Torchia, Ph.D., Timothy E.
;   REGISTRATION NUMBER: 36,700
;   REFERENCE/DOCKET NUMBER: P1086P2
; TELECOMMUNICATION INFORMATION:
;   TELEPHONE: 650/225-8674
;   TELEFAX: 650/952-9881
; INFORMATION FOR SEQ ID NO: 16:
;   SEQUENCE CHARACTERISTICS:
;     LENGTH: 664 amino acids
;     TYPE: Amino Acid
;     TOPOLOGY: Linear
;   SEQUENCE DESCRIPTION: SEQ ID NO: 16:
US-09-487-685-16

Query Match      88.0%; Score 1259; DB 4; Length 664;
Best Local Similarity 51.4%; Pred. No. 3.6e-41;
Matches 237; Conservative 11; Mismatches 14; Indels 199; Gaps 4;

QY  3 LKAFYDKV-----AEKLKE----- 16
   |:|:|:|
Db  204 LQOFFRVFSEYTYRMFLFCSCDQACAEERRQTILPSCSYEDKEKPNCLDLRGVCRDHL 263

QY  17 -----AFDWLKAFYD----- 26
   ||:|:|
Db  264 CRSRLADFHANCRASTQVTITSCPADNYQACLSYAGMIGFDMTPNVDSSTPGTIVVSPWC 323

QY  27 -----KVAEKLKEAF----- 36
   :||
Db  324 SCRGSGNMEEECEKFLRDTENPCLRNAIQAFNGTVDVNSPKGSPFQATQAPRVEKTPS 383

QY  37 -----MDKTHTC 43
   :|:|:|
Db  384 LPDDLSDSTSLGTSVITTCSTVQEQGLKANNKSKELSMCFTELNTIIPGRDPVDKTHTC 443

QY  44 PPCPAPELLGGPSVFLFPPKPKDTLMISRTPEVTCVVVDVSHEDPEVKFNWYVDGVEVHN 103
   |||
Db  444 PPCPAPELLGGPSVFLFPPKPKDTLMISRTPEVTCVVVDVSHEDPEVKFNWYVDGVEVHN 503

QY  104 AKTKPREQYNSTYRVVSVLTVLHQDWLNGKEYCKVKSNKALPAPIEKTISKAKGQPREP 163
   |||
Db  504 AKTKPREQYNSTYRVVSVLTVLHQDWLNGKEYCKVKSNKALPAPIEKTISKAKGQPREP 563

QY  164 QVYTLPPSRDELTKNOVSLTCLVKGYFSPDIAVEWESNGQPENNYKTTTPVLDSDGSFPL 223
   |||
Db  564 QVYTLPPSRDELTKNOVSLTCLVKGYFSPDIAVEWESNGQPENNYKTTTPVLDSDGSFPL 623

QY  224 YSKLTVDKSRWQGNVFCSCVMHEALHNHYTQKSLSLSPGK 264
   |||
Db  624 YSKLTVDKSRWQGNVFCSCVMHEALHNHYTQKSLSLSPGK 664

RESULT 6
US-09-487-685-18
; Sequence 18, Application US/09487685
; Patent No. 6342348
; GENERAL INFORMATION:
;   APPLICANT: Robert D. Klein, Arnon Rosenthal, Mary A. Hynes
;   TITLE OF INVENTION: Neurturin Receptor
;   NUMBER OF SEQUENCES: 19
;   CORRESPONDENCE ADDRESS:
;     ADDRESSEE: Genentech, Inc.
;     STREET: 1 DNA Way
```

```
; CITY: South San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94080
; COMPUTER READABLE FORM:
;   MEDIUM TYPE: 3.5 inch, 1.44 Mb floppy disk
;   COMPUTER: IBM PC compatible
;   OPERATING SYSTEM: PC-DOS/MS-DOS
;   SOFTWARE: WinPatin (Genentech)
; CURRENT APPLICATION DATA:
;   APPLICATION NUMBER: US/09/487,685
;   FILING DATE: 19-Jan-2000
;   CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
;   APPLICATION NUMBER: 08/957,063
;   FILING DATE: <Unknown>
;   APPLICATION NUMBER: 913
;   FILING DATE: 18-Feb-1997
; ATTORNEY/AGENT INFORMATION:
;   NAME: Torchia, Ph.D., Timothy E.
;   REGISTRATION NUMBER: 36,700
;   REFERENCE/DOCKET NUMBER: P1086P2
; TELECOMMUNICATION INFORMATION:
;   TELEPHONE: 650/225-8674
;   TELEFAX: 650/952-9881
; INFORMATION FOR SEQ ID NO: 18:
;   SEQUENCE CHARACTERISTICS:
;     LENGTH: 664 amino acids
;     TYPE: Amino Acid
;     TOPOLOGY: Linear
;   SEQUENCE DESCRIPTION: SEQ ID NO: 18:
US-09-487-685-18

Query Match      88.0%; Score 1259; DB 4; Length 664;
Best Local Similarity 51.4%; Pred. No. 3.6e-41;
Matches 237; Conservative 11; Mismatches 14; Indels 199; Gaps 4;

QY  3 LKAFYDKV-----AEKLKE----- 16
   |:|:|:|
Db  204 LQOFFRVFSEYTYRMFLFCSCDQACAEERRQTILPSCSYEDKEKPNCLDLRSLCRTDHL 263

QY  17 -----AFDWLKAFYD----- 26
   ||:|:|
Db  264 CRSRLADFHANCRASTQVTITSCPADNYQACLSYAGMIGFDMTPNVDSNPTGIVVSPWC 323

QY  27 -----KVAEKLKEAF----- 36
   :||
Db  324 NCRGSGNMEEECEKFLRDTENPCLRNAIQAFNGTVDVNSPKGSPSLPATQAPRVEKTPS 383

QY  37 -----MDKTHTC 43
   :|:|:|
Db  384 LPDDLSDSTSLGTSVITTCSTVQEQGLKANNKSKELSMCFTELNTISPGSRDPVDKTHTC 443

QY  44 PPCPAPELLGGPSVFLFPPKPKDTLMISRTPEVTCVVVDVSHEDPEVKFNWYVDGVEVHN 103
   |||
Db  444 PPCPAPELLGGPSVFLFPPKPKDTLMISRTPEVTCVVVDVSHEDPEVKFNWYVDGVEVHN 503

QY  104 AKTKPREQYNSTYRVVSVLTVLHQDWLNGKEYCKVKSNKALPAPIEKTISKAKGQPREP 163
   |||
Db  504 AKTKPREQYNSTYRVVSVLTVLHQDWLNGKEYCKVKSNKALPAPIEKTISKAKGQPREP 563

QY  164 QVYTLPPSRDELTKNOVSLTCLVKGYFSPDIAVEWESNGQPENNYKTTTPVLDSDGSFPL 223
   |||
Db  564 QVYTLPPSRDELTKNOVSLTCLVKGYFSPDIAVEWESNGQPENNYKTTTPVLDSDGSFPL 623

QY  224 YSKLTVDKSRWQGNVFCSCVMHEALHNHYTQKSLSLSPGK 264
   |||
Db  624 YSKLTVDKSRWQGNVFCSCVMHEALHNHYTQKSLSLSPGK 664

RESULT 7
US-08-802-805D-16
; Sequence 16, Application US/08802805D
```

```

; Patent No. 6372453
; GENERAL INFORMATION:
; APPLICANT: Robert D. Klein
; TITLE OF INVENTION: Neurturin Receptor
; NUMBER OF SEQUENCES: 28
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Genentech, Inc.
; STREET: 1 DNA Way
; CITY: South San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94080
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5 inch, 1.44 Mb floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: WinPatIn (Genentech)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/802,805D
; FILING DATE: 18-Feb-1997
; CLASSIFICATION: 536
; ATTORNEY/AGENT INFORMATION:
; NAME: Torchia, Phd., Timothy E.
; REGISTRATION NUMBER: 36,700
; REFERENCE/DOCKET NUMBER: P1086
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 650/225-8674
; TELEFAX: 650/952-9881
; INFORMATION FOR SEQ ID NO: 16:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 664 amino acids
; TYPE: Amino Acid
; TOPOLOGY: Linear
; US-08-802-805D-16

```

```

Query Match      88.0%; Score 1259; DB 4; Length 664;
Best Local Similarity 51.4%; Pred. No. 3.6e-41;
Matches 237; Conservative 11; Mismatches 14; Indels 199; Gaps 4;

QY  3 LKAFYDKV-----AEKLE----- 16
   |::|::|
Db  204 LRQFFDRVPSEYTYRMFCSCDQACAEARRRQITLPCSYEDKEKPNCLDLRGVCRTDHL 263
   |::|::|
QY  17 -----AFDWLKAFYD----- 26
   |::|
Db  264 CRSRLADFHANCRASTYRTITSCPADNYQACLSYAGMIGFDMTPNYVDSPTGIVVSPWC 323
   |::|
QY  27 -----KVAEKLKEAF----- 36
   |::|
Db  324 SCRGSGNMEEECEKFLRDFTEPCLRNAIQAFNGTGVNVPKGPSFQATQAPRVEKTPS 383
   |::|
QY  37 -----MDKTHTC 43
   |::|
Db  384 LPDDLSDSTSLGTSVITTCSTQEQGLKANNKSELSMCFTELNTNIIPGRDPVDKTHTC 443
   |::|
QY  44 PPCAPPELLGGSPVFLFPKPKDITLMSRTEVTCVVDVSHEDPEVKFNWYVDGVEVHN 103
   |::|
Db  444 PPCAPPELLGGSPVFLFPKPKDITLMSRTEVTCVVDVSHEDPEVKFNWYVDGVEVHN 503
   |::|
QY  104 AKTKPREQNSTYRYVSVLTVLHQDLWNGKEYCKVSNKALPAPIEKTISKAKGQPREP 163
   |::|
Db  504 AKTKPREQNSTYRYVSVLTVLHQDLWNGKEYCKVSNKALPAPIEKTISKAKGQPREP 563
   |::|
QY  164 QVYTLPPSRDELTKNQVSLTCLVKGYFSPDIAVWESNGQPENNYKTPPVLDSGDSFFL 223
   |::|
Db  564 QVYTLPPSRDEMTKNQVSLTCLVKGYFSPDIAVWESNGQPENNYKTPPVLDSGDSFFL 623
   |::|
QY  224 YSKLTVDKSRWQOGNPFSCSVMEALHNHYTQKSLSLSPCK 264
   |::|
Db  624 YSKLTVDKSRWQOGNPFSCSVMEALHNHYTQKSLSLSPCK 664

```

RESULT 8

```

US-08-802-805D-18
; Sequence 18, Application US/08802805D
; Patent No. 6372453
; GENERAL INFORMATION:
; APPLICANT: Robert D. Klein
; TITLE OF INVENTION: Neurturin Receptor
; NUMBER OF SEQUENCES: 28
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Genentech, Inc.
; STREET: 1 DNA Way
; CITY: South San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94080
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5 inch, 1.44 Mb floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: WinPatIn (Genentech)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/802,805D
; FILING DATE: 18-Feb-1997
; CLASSIFICATION: 536
; ATTORNEY/AGENT INFORMATION:
; NAME: Torchia, Phd., Timothy E.
; REGISTRATION NUMBER: 36,700
; REFERENCE/DOCKET NUMBER: P1086
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 650/225-8674
; TELEFAX: 650/952-9881
; INFORMATION FOR SEQ ID NO: 18:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 664 amino acids
; TYPE: Amino Acid
; TOPOLOGY: Linear
; US-08-802-805D-18

```

```

Query Match      88.0%; Score 1259; DB 4; Length 664;
Best Local Similarity 51.4%; Pred. No. 3.6e-41;
Matches 237; Conservative 11; Mismatches 14; Indels 199; Gaps 4;

QY  3 LKAFYDKV-----AEKLE----- 16
   |::|::|
Db  204 LRQFFDRVPSEYTYRMFCSCDQACAEARRRQITLPCSYEDKEKPNCLDLRSLCRTDHL 263
   |::|::|
QY  17 -----AFDWLKAFYD----- 26
   |::|
Db  264 CRSRLADFHANCRASTYRTITSCPADNYQACLSYAGMIGFDMTPNYVDSNPTGIVVSPWC 323
   |::|
QY  27 -----KVAEKLKEAF----- 36
   |::|
Db  324 NCRGSGNMEEECEKFLRDFTEPCLRNAIQAFNGTGVNVPKGPSLPATQAPRVEKTPS 383
   |::|
QY  37 -----MDKTHTC 43
   |::|
Db  384 LPDDLSDSTSLGTSVITTCSTQEQGLKANNKSELSMCFTELNTNISPGSRDPVDKTHTC 443
   |::|
QY  44 PPCAPPELLGGSPVFLFPKPKDITLMSRTEVTCVVDVSHEDPEVKFNWYVDGVEVHN 103
   |::|
Db  444 PPCAPPELLGGSPVFLFPKPKDITLMSRTEVTCVVDVSHEDPEVKFNWYVDGVEVHN 503
   |::|
QY  104 AKTKPREQNSTYRYVSVLTVLHQDLWNGKEYCKVSNKALPAPIEKTISKAKGQPREP 163
   |::|
Db  504 AKTKPREQNSTYRYVSVLTVLHQDLWNGKEYCKVSNKALPAPIEKTISKAKGQPREP 563
   |::|
QY  164 QVYTLPPSRDELTKNQVSLTCLVKGYFSPDIAVWESNGQPENNYKTPPVLDSGDSFFL 223
   |::|
Db  564 QVYTLPPSRDEMTKNQVSLTCLVKGYFSPDIAVWESNGQPENNYKTPPVLDSGDSFFL 623
   |::|
QY  224 YSKLTVDKSRWQOGNPFSCSVMEALHNHYTQKSLSLSPCK 264
   |::|
Db  624 YSKLTVDKSRWQOGNPFSCSVMEALHNHYTQKSLSLSPCK 664

```

RESULT 9  
US-08-378-939-10  
; Sequence 10, Application US/08378939  
; Patent No. 5876961  
; GENERAL INFORMATION:  
; APPLICANT: CROWE, JAMES SCOTT  
; APPLICANT: LEWIS, ALAN PETER  
; TITLE OF INVENTION: PRODUCTION OF ANTIBODIES  
; NUMBER OF SEQUENCES: 46  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: ROTHWELL, FIGG, ERNST & KURZ  
; STREET: 555 THIRTEENTH ST. N.W.  
; CITY: WASHINGTON  
; STATE: D. C.  
; COUNTRY: U.S.  
; ZIP: 20004  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/378,939  
; FILING DATE:  
; CLASSIFICATION: 435  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 07/952640  
; FILING DATE: 01-DEC-1992  
; ATTORNEY/AGENT INFORMATION:  
; NAME: ERNST, BARBARA G  
; REGISTRATION NUMBER: 30,377  
; REFERENCE/DOCKET NUMBER: 1808-118  
; TELEPHONE: (202) 783-6040  
; TELEFAX: (202) 783-6031  
; INFORMATION FOR SEQ ID NO: 10:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 476 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
; US-08-378-939-10

Query Match 88.0%; Score 1258; DB 2; Length 476;  
Best Local Similarity 57.8%; Pred. No. 1.8e-41;  
Matches 238; Conservative 9; Mismatches 17; Indels 148; Gaps 4;  
QY 1 DWL-----KAFYDKVAEKLKEA- 17  
Db 65 EWMGGIIPFGTPTYSQNFQGRVTITADKSTSTAHMELTSRLSEDTAVVYCATDRYQAN 124  
QY 18 FD-----WL----- 21  
Db 125 FDRARVGFDPMGQGLTVTVSSASTKGPSVFLAPSSKSTSGGTAALGCLVKDYFPPEVT 184  
QY 22 -----KAFYDKVAEKL 32  
Db 185 VSNNGALTSVHTFPVAVLQSSGLSVTVVPSSSLGTTQTYICNVNHNKPSNTKVDKRV 244  
QY 33 KEAFMDKTHTCPAPPELLGGPSVFLFPPPKKDTLMISRTPEVTCVVVDVSHEDPEVKF 92  
Db 245 EPKSCDKTHTCPAPPELLGGPSVFLFPPPKKDTLMISRTPEVTCVVVDVSHEDPEVKF 304  
QY 93 NWYVDGVEVHNAKTPREQYNSTYRVVSVLTVLHQDWLNGKEYKCKVSNKALPAPIEKT 152  
Db 305 NWYVDGVEVHNAKTPREQYNSTYRVVSVLTVLHQDWLNGKEYKCKVSNKALPAPIEKT 364  
QY 153 ISKAGQPREPQVYTLPPSRDELTKNQVSLTCLVKGFYPSDIAVEWESNGQPENNYKTP 212  
Db 365 ISKAGQPREPQVYTLPPSRDELTKNQVSLTCLVKGFYPSDIAVEWESNGQPENNYKTP 424  
QY 213 PVLDSGDSFFLYSKLTVDKSRWQGNVFSCSVHHEALHNHYTQKSLSLSPGK 264

Db 425 PVLDSGDSFFLYSKLTVDKSRWQGNVFSCSVHHEALHNHYTQKSLSLSPGK 476  
RESULT 10  
PCT-US96-10043-11  
; Sequence 11, Application PC/TUS9610043  
; GENERAL INFORMATION:  
; APPLICANT: The General Hospital Corporation  
; TITLE OF INVENTION: P-SELECTIN LIGANDS AND RELATED MOLECULES  
; NUMBER OF SEQUENCES: 14  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Fish & Richardson P.C.  
; STREET: 225 Franklin Street  
; CITY: Boston  
; STATE: MA  
; COUNTRY: USA  
; ZIP: 02210-2804  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: PCT/US96/10043  
; FILING DATE:  
; CLASSIFICATION:  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 60/000,213  
; FILING DATE: 14-JUN-1995  
; CLASSIFICATION:  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Lech, Karen F.  
; REGISTRATION NUMBER:  
; REFERENCE/DOCKET NUMBER: 00786/284001  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 617/542-5070  
; TELEFAX: 617/542-8906  
; TELEX: 200154  
; INFORMATION FOR SEQ ID NO: 11:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 437 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: not relevant  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
; PCT-US96-10043-11

Query Match 87.9%; Score 1257; DB 5; Length 437;  
Best Local Similarity 79.5%; Pred. No. 1.6e-41;  
Matches 241; Conservative 6; Mismatches 13; Indels 43; Gaps 6;  
QY 1 DW-LKAFYDK---VAEKLKEAF-----DWLKAFYDKVAEKLKEAF----- 36  
Db 139 NWGLSVADKPTTKEQLGEFEYALDCLRIKPSDVVYTDWKK---DK-CEPLEKHOKER 194  
QY 37 -----MDKTHTCPAPPELLGGPSVFLFPPPKKDTLMISRTPEVTCVV 81  
Db 195 KQEGESDPEGEKPKCDKTHTCPAPPELLGGPSVFLFPPPKKDTLMISRTPEVTCVV 254  
QY 82 DVSHEDPEVKFNWYVDGVEVHNAKTPREQYNSTYRVVSVLTVLHQDWLNGKEYKCKVS 141  
Db 255 DVSHEDPEVKFNWYVDGVEVHNAKTPREQYNSTYRVVSVLTVLHQDWLNGKEYKCKVS 314  
QY 142 NKALPAPIEKTISKAKGQPREPQVYTLPPSRDELTKNQVSLTCLVKGFYPSDIAVEWESN 201  
Db 315 NKALPAPIEKTISKAKGQPREPQVYTLPPSRDELTKNQVSLTCLVKGFYPSDIAVEWESN 374  
QY 202 QGPENNYKTTTPVLDSDGSPFLYKLTVDKSRWQGNVFSCSVHHEALHNHYTQKSLS 261  
Db 375 QGPENNYKTTTPVLDSDGSPFLYKLTVDKSRWQGNVFSCSVHHEALHNHYTQKSLS 434

```
QY 262 PGK 264
Db 435 PGK 437

RESULT 11
US-07-934-373C-22
; Sequence 22, Application US/07934373C
; Patent No. 5821337
; GENERAL INFORMATION:
; APPLICANT: Paul J. Carter
; APPLICANT: Leonard G. Presta
; TITLE OF INVENTION: Immunoglobulin Variants
; NUMBER OF SEQUENCES: 48
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Genentech, Inc.
; STREET: 1 DNA Way
; CITY: South San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94080
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5 inch, 1.44 Mb floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: WinPatin (Genentech)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/934,373C
; FILING DATE: 21-Aug-1992
; CLASSIFICATION: 530
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/US92/05126
; FILING DATE: 15-JUN-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 07/715272
; FILING DATE: 14-JUN-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: Lee, Wendy M.
; REGISTRATION NUMBER: 40,378
; REFERENCE/DOCKET NUMBER: P0709P2
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 650/225-1994
; TELEFAX: 650/952-9881
; INFORMATION FOR SEQ ID NO: 22:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 454 amino acids
; TYPE: Amino Acid
; TOPOLOGY: Linear
US-07-934-373C-22

Query Match 87.9%; Score 1257; DB 2; Length 454;
Best Local Similarity 54.9%; Pred. No. 1.7e-41;
Matches 234; Conservative 11; Mismatches 11; Indels 170; Gaps 3;

QY 2 WLKAFYDKVAEKLKEAFDWLAKF----- 24
Db 36 WMKQSHGK-----SLEWIGGFNPKNGSSHNORFMDKATLAVDKSTSTAYMELRSLTS 88

QY 25 -----YD----- 26
Db 89 EDGIYYCARWRGLNGYDFVYEDVWGAGTTVTVSSASTKGPSVFLPAPSSKSTSGGTAA 148

QY 27 ----- 26
Db 149 LGCLVLDYFPEPVTVSNWNSGALTSVHTFPVAVLQSSGLYSLSSVTVTPSSSLGTQTICN 208

QY 27 -----KVAELKEAFMDKTHTCPCPAPELLGGPSVFLFPKPKDMLMISRTPEVTC 78
Db 209 VNHKPSNTKVDKKVEPKSCDKTHTCPCPAPELLGGPSVFLFPKPKDMLMISRTPEVTC 268

QY 79 VVDVSHEDPEVKFNWTVDGVEVHNATKPREQYNSTYRVSVLTVLHODWLNKGEYKC 138
Db 269 VVDVSHEDPEVKFNWTVDGVEVHNATKPREQYNSTYRVSVLTVLHODWLNKGEYKC 328

QY 139 KVSNKALPAPIEKTISKAKGPQVYTLPPSRDELTKNOVSLTCLVKGFYPSDIAVEW 198
Db 329 KVSNKALPAPIEKTISKAKGPQVYTLPPSRDELTKNOVSLTCLVKGFYPSDIAVEW 388

QY 199 ESNQGPENNYKTTTPVLDSDGSFFLYSKLTVDKSRWQGNVFCVSMHEALHNNHYTKSL 258
Db 389 ESNQGPENNYKTTTPVLDSDGSFFLYSKLTVDKSRWQGNVFCVSMHEALHNNHYTKSL 448

QY 259 SLSPGK 264
Db 449 SLSPGK 454

RESULT 12
US-08-437-642B-22
; Sequence 22, Application US/08437642B
; Patent No. 6054297
; GENERAL INFORMATION:
; APPLICANT: Paul J. Carter
; APPLICANT: Leonard G. Presta
; TITLE OF INVENTION: Immunoglobulin Variants
; NUMBER OF SEQUENCES: 47
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Genentech, Inc.
; STREET: 1 DNA Way
; CITY: South San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94080
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5 inch, 1.44 Mb floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: WinPatin (Genentech)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/437,642B
; FILING DATE: 09-May-1995
; CLASSIFICATION: 530
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 07/934373
; FILING DATE: 21-AUG-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/146206
; FILING DATE: 17-NOV-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/US92/05126
; FILING DATE: 15-JUN-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 07/715272
; FILING DATE: 14-JUN-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: Lee, Wendy M.
; REGISTRATION NUMBER: 40,378
; REFERENCE/DOCKET NUMBER: P0709P2C1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 650/225-1994
; TELEFAX: 650/952-9881
; INFORMATION FOR SEQ ID NO: 22:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 454 amino acids
; TYPE: Amino Acid
; TOPOLOGY: Linear
US-08-437-642B-22

Query Match 87.9%; Score 1257; DB 3; Length 454;
Best Local Similarity 54.9%; Pred. No. 1.7e-41;
Matches 234; Conservative 11; Mismatches 11; Indels 170; Gaps 3;

QY 2 WLKAFYDKVAEKLKEAFDWLAKF----- 24
Db 36 WMKQSHGK-----SLEWIGGFNPKNGSSHNORFMDKATLAVDKSTSTAYMELRSLTS 88
```

QY 25 -----YD----- 26  
Db 89 EDGIYICARWGLNYGDFVRYDFVWAGCTTVTVSSASTKGPSPVFLPAPSSKSTSGGTAA 148  
QY 27 ----- 26  
Db 149 LGCLVKDYFPEPVTVSWNSGALTSGVHTTFAVLQSSGLYSLSSVTVTPSSSLGTQTYICN 208  
QY 27 -----KVAEKLKEAFMDKTHTCPCPAPELLGGPSVFLFPPPKDFTLMISRTPPEVTC 78  
Db 209 VNHKPSNTKVDKVEPKSCDKTHTCPCPAPELLGGPSVFLFPPPKDFTLMISRTPPEVTC 268  
QY 79 VYDVSHEDPEVKFNMYVDGVEVHNNAKTKPREQYNSTYRVYVSVLTVLHODMNLNGKEYKC 138  
Db 269 VYDVSHEDPEVKFNMYVDGVEVHNNAKTKPREQYNSTYRVYVSVLTVLHODMNLNGKEYKC 328  
QY 139 KVSNAKALPAPIEKTISKAKGQPREPOVYITLPPSRDELTKNQVSLTCLVKGFYPSDIAVEW 198  
Db 329 KVSNAKALPAPIEKTISKAKGQPREPOVYITLPPSRDELTKNQVSLTCLVKGFYPSDIAVEW 388  
QY 199 ESNQGPENNYKTPPPVLDSDGSFFLYSKLTVDKSRWQOGNVFSCSVMHALHNNHYTQKSL 258  
Db 389 ESNQGPENNYKTPPPVLDSDGSFFLYSKLTVDKSRWQOGNVFSCSVMHALHNNHYTQKSL 448  
QY 259 SLSPGK 264  
Db 449 SLSPGK 454

RESULT 13  
US-08-146-206C-22  
; Sequence 22, Application US/08146206C  
; Patent No. 6407213  
; GENERAL INFORMATION:  
; APPLICANT: Carter, Paul J.  
; APPLICANT: Presta, Leonard G.  
; TITLE OF INVENTION: Method for Making Humanized Antibodies  
; NUMBER OF SEQUENCES: 26  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Genentech, Inc.  
; STREET: 1 DNA Way  
; CITY: South San Francisco  
; STATE: California  
; COUNTRY: USA  
; ZIP: 94080  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: 3.5 inch, 1.44 Mb floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: WinPatIn (Genentech)  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/146,206C  
; FILING DATE: 17-No. 6407213-1993  
; CLASSIFICATION: 530  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 07/715272  
; FILING DATE: 14-JUN-1991  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Lee, Wendy M.  
; REGISTRATION NUMBER: 40,378  
; REFERENCE/DOCKET NUMBER: P0709P1  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 650/225-1994  
; TELEFAX: 650/952-9881  
; INFORMATION FOR SEQ ID NO: 22:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 454 amino acids  
; TYPE: Amino Acid  
; TOPOLOGY: Linear  
US-08-146-206C-22

Query Match 87.9%; Score 1257; DB 4; Length 454;  
Best Local Similarity 54.9%; Pred. No. 1.7e-41;

Matches 234; Conservative 11; Mismatches 11; Indels 170; Gaps 3;  
QY 2 WLKAFDYKVAEKLKEAFMDLKAF----- 24  
Db 36 WNKQSHGK-----SLEWIGGFNPKNGGSSHNQRPMDKATLAVDKSTSTAYMELRSLTS 88  
QY 25 -----YD----- 26  
Db 89 EDGIYICARWGLNYGDFVRYDFVWAGCTTVTVSSASTKGPSPVFLPAPSSKSTSGGTAA 148  
QY 27 ----- 26  
Db 149 LGCLVKDYFPEPVTVSWNSGALTSGVHTTFAVLQSSGLYSLSSVTVTPSSSLGTQTYICN 208  
QY 27 -----KVAEKLKEAFMDKTHTCPCPAPELLGGPSVFLFPPPKDFTLMISRTPPEVTC 78  
Db 209 VNHKPSNTKVDKVEPKSCDKTHTCPCPAPELLGGPSVFLFPPPKDFTLMISRTPPEVTC 268  
QY 79 VYDVSHEDPEVKFNMYVDGVEVHNNAKTKPREQYNSTYRVYVSVLTVLHODMNLNGKEYKC 138  
Db 269 VYDVSHEDPEVKFNMYVDGVEVHNNAKTKPREQYNSTYRVYVSVLTVLHODMNLNGKEYKC 328  
QY 139 KVSNAKALPAPIEKTISKAKGQPREPOVYITLPPSRDELTKNQVSLTCLVKGFYPSDIAVEW 198  
Db 329 KVSNAKALPAPIEKTISKAKGQPREPOVYITLPPSRDELTKNQVSLTCLVKGFYPSDIAVEW 388  
QY 199 ESNQGPENNYKTPPPVLDSDGSFFLYSKLTVDKSRWQOGNVFSCSVMHALHNNHYTQKSL 258  
Db 389 ESNQGPENNYKTPPPVLDSDGSFFLYSKLTVDKSRWQOGNVFSCSVMHALHNNHYTQKSL 448  
QY 259 SLSPGK 264  
Db 449 SLSPGK 454

RESULT 14  
PCT-US93-07832-22  
; Sequence 22, Application PC/TUS9307832  
; GENERAL INFORMATION:  
; APPLICANT: Genentech, Inc.  
; TITLE OF INVENTION: Immunoglobulin Variants  
; NUMBER OF SEQUENCES: 40  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Genentech, Inc.  
; STREET: 460 Point San Bruno Blvd  
; CITY: South San Francisco  
; STATE: California  
; COUNTRY: USA  
; ZIP: 94080  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: 5.25 inch, 360 Kb floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: patin (Genentech)  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: PCT/US93/07832  
; FILING DATE: 19930820  
; CLASSIFICATION:  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 07/715272  
; FILING DATE: 14-JUN-1991  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: PCT/US92/05126  
; FILING DATE: 15-JUN-1992  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 07/934373  
; FILING DATE: 21-AUG-1992  
; ATTORNEY/AGENT INFORMATION:  
; NAME:  
; REGISTRATION NUMBER:  
; REFERENCE/DOCKET NUMBER: 709P2PCT  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE:

Search completed: April 21, 2003, 10:50:25  
Job time : 16.7529 secs





GenCore version 5.1.4\_p5\_4578  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: April 21, 2003, 10:43:21 ; Search time 12.1647 Seconds  
(without alignments)  
1640.982 Million cell updates/sec

Title: 7LINK7LINK2

Perfect score: 1430

Sequence: 1 DWLAFYDKVAEKLKEAFDW.....MHEALHNYTKSLSPGK 264

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.0

Searched: 288829 seqs, 75613885 residues

Total number of hits satisfying chosen parameters: 288829

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published\_Applications\_AA.\*

- 1: /cgn2\_6/ptodata/2/pubpaa/US08\_NEW\_PUB.pep.\*
- 2: /cgn2\_6/ptodata/2/pubpaa/PT\_NEW\_PUB.pep.\*
- 3: /cgn2\_6/ptodata/2/pubpaa/US06\_NEW\_PUB.pep.\*
- 4: /cgn2\_6/ptodata/2/pubpaa/US06\_PUBCOMB.pep.\*
- 5: /cgn2\_6/ptodata/2/pubpaa/US07\_NEW\_PUB.pep.\*
- 6: /cgn2\_6/ptodata/2/pubpaa/US07\_PUBCOMB.pep.\*
- 7: /cgn2\_6/ptodata/2/pubpaa/PTUS\_PUBCOMB.pep.\*
- 8: /cgn2\_6/ptodata/2/pubpaa/US08\_PUBCOMB.pep.\*
- 9: /cgn2\_6/ptodata/2/pubpaa/US09\_NEW\_PUB.pep.\*
- 10: /cgn2\_6/ptodata/2/pubpaa/US09\_PUBCOMB.pep.\*
- 11: /cgn2\_6/ptodata/2/pubpaa/US10\_NEW\_PUB.pep.\*
- 12: /cgn2\_6/ptodata/2/pubpaa/US10\_PUBCOMB.pep.\*
- 13: /cgn2\_6/ptodata/2/pubpaa/US60\_NEW\_PUB.pep.\*
- 14: /cgn2\_6/ptodata/2/pubpaa/US60\_PUBCOMB.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1268	88.7	594	10 US-09-815-108-22	Sequence 22, Appl
2	1268	88.7	951	9 US-09-935-868-9	Sequence 9, Appl
3	1268	88.7	1158	9 US-09-935-868-26	Sequence 26, Appl
4	1268	88.7	1168	9 US-09-935-868-24	Sequence 24, Appl
5	1263	88.3	567	10 US-09-825-561A-16	Sequence 16, Appl
6	1263	88.3	859	9 US-09-935-868-7	Sequence 7, Appl
7	1261	88.2	476	9 US-10-124-905-12	Sequence 12, Appl
8	1261	88.2	476	9 US-09-948-429B-12	Sequence 12, Appl
9	1260	88.1	399	9 US-09-832-659-2	Sequence 2, Appl
10	1259	88.0	547	10 US-09-746-359A-54	Sequence 54, Appl
11	1259	88.0	571	10 US-09-746-359A-53	Sequence 53, Appl
12	1258	88.0	475	10 US-09-740-002-27	Sequence 27, Appl
13	1256	87.8	347	9 US-10-091-236-17	Sequence 17, Appl
14	1256	87.8	347	9 US-10-091-313-7	Sequence 7, Appl
15	1256	87.8	347	9 US-10-091-268-7	Sequence 7, Appl
16	1256	87.8	347	10 US-09-796-033-8	Sequence 8, Appl
17	1256	87.8	347	10 US-09-730-465-8	Sequence 8, Appl
18	1256	87.8	776	9 US-09-935-868-36	Sequence 36, Appl
19	1256	87.8	776	9 US-09-935-868-40	Sequence 40, Appl

20	1256	87.8	776	9 US-09-935-868-44	Sequence 44, Appl
21	1256	87.8	780	9 US-09-935-868-34	Sequence 34, Appl
22	1256	87.8	780	9 US-09-935-868-38	Sequence 38, Appl
23	1256	87.8	780	9 US-09-935-868-42	Sequence 42, Appl
24	1256	87.8	910	9 US-09-935-868-28	Sequence 28, Appl
25	1255	87.8	423	9 US-09-832-659-44	Sequence 44, Appl
26	1255	87.8	452	9 US-09-726-258-71	Sequence 71, Appl
27	1255	87.8	475	10 US-09-740-002-25	Sequence 25, Appl
28	1254	87.7	592	9 US-09-935-868-8	Sequence 8, Appl
29	1254	87.7	691	9 US-09-935-868-20	Sequence 20, Appl
30	1254	87.7	694	9 US-09-935-868-22	Sequence 22, Appl
31	1254	87.7	1194	9 US-10-191-029-10	Sequence 10, Appl
32	1253	87.6	418	9 US-09-832-659-42	Sequence 42, Appl
33	1253	87.6	731	10 US-09-825-012-46	Sequence 46, Appl
34	1253	87.6	741	10 US-09-825-012-55	Sequence 55, Appl
35	1252	87.6	450	9 US-09-996-288-228	Sequence 228, App
36	1252	87.6	450	10 US-09-796-848A-47	Sequence 47, Appl
37	1252	87.6	482	10 US-09-824-286-2	Sequence 2, Appl
38	1251	87.5	450	9 US-09-996-288-216	Sequence 216, App
39	1251	87.5	450	9 US-09-996-288-218	Sequence 218, App
40	1251	87.5	450	9 US-09-996-288-250	Sequence 250, App
41	1251	87.5	450	10 US-09-796-848A-43	Sequence 43, Appl
42	1251	87.5	450	10 US-09-796-848A-49	Sequence 49, Appl
43	1251	87.5	451	10 US-09-822-698A-26	Sequence 26, Appl
44	1251	87.5	453	9 US-09-925-179-8	Sequence 8, Appl
45	1251	87.5	453	10 US-09-802-077-8	Sequence 8, Appl

ALIGNMENTS

RESULT 1

US-09-815-108-22  
; Sequence 22, Application US/09815108  
; Patent No. US20020009776A1  
; GENERAL INFORMATION:  
; APPLICANT: Saris, Christiaan M.  
; APPLICANT: Sharon, Mu X.  
; APPLICANT: Xia, Min  
; APPLICANT: Boone, Thomas Charles  
; APPLICANT: Covey, Todd  
; TITLE OF INVENTION: Fibroblast Growth Factor Receptor-Like Molecules and  
; FILE OF INVENTION: Uses Thereof  
; FILE REFERENCE: 99-513-A  
; CURRENT APPLICATION NUMBER: US/09/815,108  
; PRIOR FILING DATE: 2001-03-22  
; PRIOR APPLICATION NUMBER: 60/191,379  
; PRIOR FILING DATE: 2000-03-22  
; NUMBER OF SEQ ID NOS: 22  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 22  
; LENGTH: 594  
; TYPE: PRT  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Description of Artificial Sequence: murine FGFR-L  
; OTHER INFORMATION: extracellular domain-Fc fusion polypeptide  
US-09-815-108-22

Query Match	88.7%	Score 1268;	DB 10;	Length 594;
Best Local Similarity	57.6%	Pred. No. 3.3e-32;		
Matches 239;	Conservative	9;	Mismatches 15;	Indels 152; Gaps 4;
QY	2	WLK-----		4
Db	180	WMKDDQTLHLEASHKHKWTLSLKNLPEDSGKTCRVSNKAGAINATYKVDVIQTR		239
QY	5	-----AFYDKVAEKLKEAFDWLK-----		22
Db	240	SRPVLGTGTHPVNTVDFGGTTSFQCKVRSDVKPVIQWLKRVYSGRHNSTIDVGQKF		299
QY	23	-----AFYDK-----		37

```
Db 300 VVLPTGDVMSRPGDGYLNKLLISRARQDDAGMYICLGANTMGYSFSAFLTVLPDPKPPG 359
Qy 38 -----DKTHTCPAPPELLGGPSVFLFPPKPKDTLMISRTPEVTCVVVDVSHEDPE 89
Db 360 PPMASSSVDKTHTCPCPCAPPELLGGPSVFLFPPKPKDTLMISRTPEVTCVVVDVSHEDPE 419
Qy 90 VKFNQYVGVGVHNAKTKPREQYNSTYRVVSVLTVLHQDWLNGKEYKCKVSNKALPAPI 149
Db 420 VKFNQYVGVGVHNAKTKPREQYNSTYRVVSVLTVLHQDWLNGKEYKCKVSNKALPAPI 479
Qy 150 EKTISKAKGQPREPOVYTLPPSRDELTKNQVSLTCLVKGFYPSDIAVEWESNGQPENNYK 209
Db 480 EKTISKAKGQPREPOVYTLPPSRDELTKNQVSLTCLVKGFYPSDIAVEWESNGQPENNYK 539
Qy 210 TTPPVLDSDGSFFLYSLKLTVDKSRWQOGNVFSCSVMHHEALHNHYTQKSLSLSPGK 264
Db 540 TTPPVLDSDGSFFLYSLKLTVDKSRWQOGNVFSCSVMHHEALHNHYTQKSLSLSPGK 594

RESULT 2
US-09-935-868-9
; Sequence 9, Application US/09935868
; Patent No. US20020164690A1
; GENERAL INFORMATION:
; APPLICANT: Regeneron Pharmaceuticals, Inc
; TITLE OF INVENTION: Receptor Based Antagonists, and Methods of Making and Using
; FILE REFERENCE: REG 203D
; CURRENT APPLICATION NUMBER: US/09/935,868
; PRIOR FILING DATE: 2002-04-11
; PRIOR APPLICATION NUMBER: PCT/US99/22045
; PRIOR FILING DATE: 1999-09-22
; NUMBER OF SEQ ID NOS: 52
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 9
; LENGTH: 951
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-935-868-9

Query Match 88.7%; Score 1268; DB 9; Length 951;
Best Local Similarity 38.4%; Pred. No. 9.7e-32;
Matches 246; Conservative 8; Mismatches 10; Indels 376; Gaps 7;

Qy 1 DW----- 2
Db 312 DWSEASGITYEDRPSKAPSFYKIDPSHTQGYRTVQLVWKTLPPFEANGKILDYEVTLT 371
Qy 3 -----LK 4
Db 372 RWKSHLQNTVNAKLTVNLNDRYLATLTVRNLVGKSDAAVLTIPACDFQATHPVMDLK 431
Qy 5 AF-----YDK-----VAEK-- 13
Db 432 AFPKDNMLVWETTPRESVKYKYLEWCVLSDKAPCITDQOEDGTVHRTYLRGNLAESKC. 491
Qy 14 -----LKEA-----FDW----- 20
Db 492 YLITVTPYADGPGSPESIKAYLKQAPPSKGTVRTKKVGNKNEAVLEWDLQLPVDVQNGFI 551
Qy 21 -----LKAFLD----- 26
Db 552 RNYTIFRTIIGNETAVNVDSSHTYTLSSLTSDTYMVRMAAYTDEGGKDGPEFTTTP 611
Qy 27 ----- 26
Db 612 KFAQGEIESGASTKGPSVFPLAPSSKSTSGGTAALGCLVKDYFPEPTVTSWNSGALTSGV 671
Qy 27 -----KVAEKLKEAFMDKTHTCP 44
Db 672 HTEPAVLQSSGLYSLSVVTPSSSLGQTQYICNVNHNKPSNTKVDKKVEPKSCDKTHTCP 731
Qy 45 PCPAPPELLGGPSVFLFPPKPKDTLMISRTPEVTCVVVDVSHEDPEVKFNWYVDGVEVHNA 104
|||||
```

```
Db 732 PCPAPPELLGGPSVFLFPPKPKDTLMISRTPEVTCVVVDVSHEDPEVKFNWYVDGVEVHNA 791
Qy 105 KTKPREQYNSTYRVVSVLTVLHQDWLNGKEYKCKVSNKALPAPIEKTIISKAKGQPREPQ 164
Db 792 KTKPREQYNSTYRVVSVLTVLHQDWLNGKEYKCKVSNKALPAPIEKTIISKAKGQPREPQ 851
Qy 165 VYTLPPSRDELTKNQVSLTCLVKGFYPSDIAVEWESNGQPENNYKTTTPPVLDSDGSFFLY 224
Db 852 VYTLPPSRDELTKNQVSLTCLVKGFYPSDIAVEWESNGQPENNYKTTTPPVLDSDGSFFLY 911
Qy 225 SKLTVDKSRWQOGNVFSCSVMHHEALHNHYTQKSLSLSPGK 264
Db 912 SKLTVDKSRWQOGNVFSCSVMHHEALHNHYTQKSLSLSPGK 951

RESULT 3
US-09-935-868-26
; Sequence 26, Application US/09935868
; Patent No. US20020164690A1
; GENERAL INFORMATION:
; APPLICANT: Regeneron Pharmaceuticals, Inc
; TITLE OF INVENTION: Receptor Based Antagonists, and Methods of Making and Using
; FILE REFERENCE: REG 203D
; CURRENT APPLICATION NUMBER: US/09/935,868
; PRIOR FILING DATE: 2002-04-11
; PRIOR APPLICATION NUMBER: PCT/US99/22045
; PRIOR FILING DATE: 1999-09-22
; NUMBER OF SEQ ID NOS: 52
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 26
; LENGTH: 1158
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-935-868-26

Query Match 88.7%; Score 1268; DB 9; Length 1158;
Best Local Similarity 37.1%; Pred. No. 1.5e-31;
Matches 242; Conservative 12; Mismatches 9; Indels 390; Gaps 6;

Qy 2 WLKA-----FVDKV-----AEKL----- 14
Db 506 WVEAENALGKVTSDHINFDVPYKVPKPNPPHNLVINSEELSLKLTWNPISIKSVILK 565
Qy 15 -----KEAFDW----- 20
Db 566 YNIQYRTKDASTWQIPEPTASTRSSTVQDLAPFTYEVFIRCMKEDGKGYSDNSEE 625
Qy 21 ----- 20
Db 626 ASGITIEDRPSKAPSFYKIDPSHTQGYRTVQLVWKTLPPFEANGKILDYEVTLTRMKSH 685
Qy 21 -----LKAFLD----- 24
Db 686 LQNTVNAKLTVNLNDRYLATLTVRNLVGKSDAAVLTIPACDFQATHPVMDLKAPFKD 745
Qy 25 ----- 24
Db 746 NMLVWETTPRESVKYKYLEWCVLSDKAPCITDQOEDGTVHRTYLRGNLAESKYLITV 805
Qy 25 -----YDKVAEKLKEAFM----- 37
Db 806 TPVYADGPGSPESIKAYLKQAPPSKGTVRTKKVGNKNEAVLEWDLQLPVDVQNGFI 865
Qy 38 ----- 37
Db 866 FYRTIIGNETAVNVDSSHTYTLSSLTSDTYMVRMAAYTDEGGKDGPEFTTTPKFAQG 925
Qy 38 -----DKTHTCPAPPELLGGPSVFLFPPKPKDTLMISRTPEVTCVVVDVSHEDPEVK 91
Db 926 EIESGGDKTHTCPCPCAPPELLGGPSVFLFPPKPKDTLMISRTPEVTCVVVDVSHEDPEVK 985
Qy 92 FNVYVDGVEVHNAKTKPREQYNSTYRVVSVLTVLHQDWLNGKEYKCKVSNKALPAPIEK 151
|||||
```



## RESULT 6

```

US-09-935-868-7
; Sequence 7, Application US/09935868
; Patent No. US20020164690A1
; GENERAL INFORMATION:
; APPLICANT: Regeneron Pharmaceuticals, Inc
; TITLE OF INVENTION: Receptor Based Antagonists, and Methods of Making and Using
; FILE REFERENCE: REG 203D
; CURRENT APPLICATION NUMBER: US/09/935,868
; CURRENT FILING DATE: 2002-04-11
; PRIOR APPLICATION NUMBER: PCT/US99/22045
; PRIOR FILING DATE: 1999-09-22
; NUMBER OF SEQ ID NOS: 52
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 7
; LENGTH: 859
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-935-868-7

```

Query Match	88.3%	Score 1263;	DB 9;	Length 859;
Best Local Similarity	44.3%;	Pred. No. 1.1e-31;		
Matches 240;	Conservative	12;	Mismatches 12;	Indels 278;
Gaps				

QY	1	DW	-----	2
Db	312	DWSEASGITYEDRPSKAPSWYKIDPSHTQGYRTVOLVWMTLPPPEANGKILDYEWTLT	371	
QY	3		-----LK 4	
Db	372	RWKSHLQNYTNATKTLTVNLTDNDRYLATLTVRNLVGKSDAAVLITIPACDFQATHPVMDLK	431	
QY	5	AF	-----YDKVAEKLKEAFDW	20
Db	432	AFPKDNMLWVWTTTPRESVKYKYLEWCVLSDKAPCITDWOQEDGTVHRTYLRGNLAESKC	491	
QY	21		-----LKA	
Db	492	YLITVTPVYADGPGSPESIKAYLKQAPSGKPTVRTKVGKNEAVLEWDQLPVDVONGFI	551	
QY	38		-----	37
Db	552	RNTYIFRTIIGNETAVNDSSHTEYTLSSLTSDTLVYVMAAYTDGKGKDGEFFTTTP	611	
QY	38		-----DKTHTPCCPAPELGGPSVFLFPPKPKDTLMIKRTPEVTCVVD	82
Db	612	KFAQGEIESGEPKSCDKTHTCPAPELGGPSVFLFPPKPKDTLMIKRTPEVTCVVD	671	
QY	83	VSHEDVEVFNWYVDGVEVHNAKTKPREQYNSTYRVVSVLTVLHQDWLNGKEYKCKVSN	142	
Db	672	VSHEDPEVAFNMYVDGVEVHNAKTKPREQYNSTYRVVSVLTVLHQDWLNGKEYKCKVSN	731	
QY	143	KALPAPIETKISKAKOPREPQVYITLPPSRDELTKNOVSLTCLVKGFGYPSDIAVEWESNG	202	
Db	732	KALPAPIETKISKAKOPREPQVYITLPPSRDELTKNOVSLTCLVKGFGYPSDIAVEWESNG	791	
QY	203	QPENNYKTTTPVLDSDGSEFFLYSKLTVDKSRWQOGNVFSCSVHMEALHNNHYTKQSLSLSP	262	
Db	792	QPENNYKTTTPVLDSDGSEFFLYSKLTVDKSRWQOGNVFSCSVHMEALHNNHYTKQSLSLSP	851	
QY	263	GK	264	
Db	852	GK	853	

## RESULT 7

```

RESULTS
US-10-124-905-12
; Sequence 12, Application US/10124905
; Patent No. US20020166136A1
; GENERAL INFORMATION:
; APPLICANT: Anderson, Darrell R.
; TITLE OF INVENTION: "MONKEY MONOCLONAL ANTIBODIES SPECIFIC
; TITLE OF INVENTION: TO HUMAN B7.1 AND/OR B7.2 PRIMATIZED FORMS
; THEREOF.

```

Db 452 FSCVMHEALHNHYTKQSLSPGK 476

RESULT 8

US-09-948-429B-12

Sequence 12, Application US/09948429B

Patent No. US20020177689A1

GENERAL INFORMATION:

APPLICANT: Anderson, Darrell R.

TITLE OF INVENTION: "MONKEY MONOCLONAL ANTIBODIES SPECIFIC TO HUMAN B7.1 AND/OR B7.2 PRIMATIZED FORMS THEREOF,

TITLE OF INVENTION: PHARMACEUTICAL COMPOSITIONS CONTAINING, AND USE THEREOF AS

TITLE OF INVENTION: IMMUNOSUPPRESSANTS"

NUMBER OF SEQUENCES: 12

CORRESPONDENCE ADDRESS:

ADDRESSEE: BURNS, DOANE, SWECKER & MATHIS

STREET: 699 Prince Street

CITY: Alexandria

STATE: VA

COUNTRY: USA

ZIP: 22314

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

OPERATING SYSTEM: IBM PC compatible

SOFTWARE: PatentIn Release #1.0, Version #1.30

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/09/948,429B

FILING DATE:

CLASSIFICATION:

PRIOR APPLICATION DATA:

APPLICATION NUMBER: 09/383,916

FILING DATE:

APPLICATION NUMBER: US 08/487,550

FILING DATE: 07-JUN-1995

ATTORNEY/AGENT INFORMATION:

NAME: Teskin, Robin L.

REGISTRATION NUMBER: 35,030

REFERENCE/DOCKET NUMBER: 012712-131

TELECOMMUNICATION INFORMATION:

TELEPHONE: 703-836-6620

TELEFAX: 703-836-2021

INFORMATION FOR SEQ ID NO: 12:

SEQUENCE CHARACTERISTICS:

LENGTH: 476 amino acids

TYPE: amino acid

TOPOLOGY: linear

MOLECULE TYPE: protein

US-09-948-429B-12

Query Match 88.2%; Score 1261; DB 9; Length 476;

Best Local Similarity 53.0%; Pred. No. 3.3e-32;

Matches 236; Conservative 7; Mismatches 13; Indels 189; Gaps 2;

Qy 9 KVAEKLK-----EAFDWLKAFYD----- 26

Db 32 KPSETLSLTCVAGSGSISGGYGWGIROPPGKGLWIGSFYSSGNTYINPSLKSQVITIS 91

Qy 27 ----- 26

Db 92 TDTSKNOFSLKNSMTAADTAVYYCVDRRLFVSGVMYNNWFDVWGPGLVTVTSASTKG 151

Qy 27 ----- 26

Db 152 PSVFPLAPSSKSTSGGTAALGCLVKDYFPEPTVYSWNSGALTSGVHTFPAVLQSSGLYSL 211

Qy 27 -----KVAEKLKAFMDKTHCTCPCPAPPELLGGPSVFL 59

Db 212 SSVTVTPSSSLGTQTYICNVNHKPSNTKVDKKAEPKSCDKTHCTCPAPPELLGGPSVFL 271

Qy 60 FPPKPKDTLMISRTPEVTCVVDVSHEDPEVKFNWYVDGVEVHNAKTKPREEQYNSTYRV 119

Db 272 FPPKPKDTLMISRTPEVTCVVDVSHEDPEVKFNWYVDGVEVHNAKTKPREEQYNSTYRV 331

Qy 120 VSVLTVLHODWLNCKEYKCKVSNKALPAPIEKTISKAKGQPREQVYTLPPSRDELTKNQ 179

Db 332 VSVLTVLHODWLNCKEYKCKVSNKALPAPIEKTISKAKGQPREQVYTLPPSRDELTKNQ 391

Qy 180 VSLTCLVKGFYPSDIAVEWESNGQPENNYKTPPVLDSDGSGFFLYSKLTVDKSRWQQGNV 239

Db 392 VSLTCLVKGFYPSDIAVEWESNGQPENNYKTPPVLDSDGSGFFLYSKLTVDKSRWQQGNV 451

Qy 240 FSCSVMEALHNHYTKQSLSPGK 264

Db 452 FSCSVMEALHNHYTKQSLSPGK 476

RESULT 9

US-09-832-659-2

Sequence 2, Application US/09832659

Patent No. US20020155547A1

GENERAL INFORMATION:

APPLICANT: BIOGEN, INC.

TITLE OF INVENTION: Interferon-Beta Fusion Proteins and Uses

FILE REFERENCE: A064PCTSEQ

CURRENT APPLICATION NUMBER: US/09/832,659

CURRENT FILING DATE: 2001-04-11

PRIOR APPLICATION NUMBER: 60/120,237

PRIOR FILING DATE: 1999-02-16

PRIOR APPLICATION NUMBER: 60/104,491

PRIOR FILING DATE: 1998-10-16

NUMBER OF SEQ ID NOS: 44

SOFTWARE: PatentIn Ver. 2.0

SEQ ID NO 2

LENGTH: 399

TYPE: PRT

ORGANISM: murine

US-09-832-659-2

Query Match 88.1%; Score 1260; DB 9; Length 399;

Best Local Similarity 82.9%; Pred. No. 2.3e-32;

Matches 238; Conservative 5; Mismatches 10; Indels 34; Gaps 3;

Qy 3 LKAFYDKVAEKLKE-----AFDWLKAFY-----DKVAEKLKEAFM 37

Db 122 LKRYIGRILHYLKAKEYSHCAWTIVRVEILRNFFYINRLTGYLRNDDDDKV----- 172

Qy 38 DKHTCPCPAPPELLGGPSVFLPPKPKDTLMISRTPEVTCVVDVSHEDPEVKFNWYVD 97

Db 173 DKHTCPCPAPPELLGGPSVFLPPKPKDTLMISRTPEVTCVVDVSHEDPEVKFNWYVD 232

Qy 98 GVEVHNAKTKPREQYNSTYRVVSVLTVLHODWLNCKEYKCKVSNKALPAPIEKTISKAK 157

Db 233 GVEVHNAKTKPREQYNSTYRVVSVLTVLHODWLNCKEYKCKVSNKALPAPIEKTISKAK 292

Qy 158 GQPREQVYTLPPSRDELTKNQVSLTCLVKGFYPSDIAVEWESNGQPENNYKTPPVLDSD 217

Db 293 GQPREQVYTLPPSRDELTKNQVSLTCLVKGFYPSDIAVEWESNGQPENNYKTPPVLDSD 352

Qy 218 DGSFFLYSKLTVDKSRWQQGNVFCFSVMHEALHNHYTKQSLSPGK 264

Db 353 DGSFFLYSKLTVDKSRWQQGNVFCFSVMHEALHNHYTKQSLSPGK 399

RESULT 10

US-09-746-359A-54

Sequence 54, Application US/09746359A

Patent No. US20020042366A1

GENERAL INFORMATION:

APPLICANT: Thompson, Penny

APPLICANT: Foster, Donald C.

APPLICANT: Xu, Wenfeng

APPLICANT: Madden, Karen L.

APPLICANT: Kelly, James D.

APPLICANT: Sprecher, Cindy A.

APPLICANT: Blumberg, Hal

; APPLICANT: Eagan, Maribeth A.  
; APPLICANT: Jaspers, Stephen R.  
; APPLICANT: Chandrasekher, Yasmin A.  
; APPLICANT: No. US20020042366Alak, Julia E.  
; TITLE OF INVENTION: Method for Treating Inflammation  
; FILE REFERENCE: 99-108  
; CURRENT APPLICATION NUMBER: US/09/746,359A  
; CURRENT FILING DATE: 2001-05-21  
; PRIOR APPLICATION NUMBER: 60/171,969  
; PRIOR FILING DATE: 1999-12-23  
; PRIOR APPLICATION NUMBER: 60/213,341  
; PRIOR FILING DATE: 2000-06-22  
; NUMBER OF SEQ ID NOS: 72  
; SOFTWARE: FastSeq for Windows Version 3.0  
; SEQ ID NO 54  
; LENGTH: 547  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-09-746-359A-54

Query Match 88.0%; Score 1259; DB 10; Length 547;  
Best Local Similarity 48.6%; Pred. No. 5.2e-32;  
Matches 241; Conservative 4; Mismatches 18; Indels 233; Gaps 4;  
QY 2 WLK-----AFYDKV-----10  
Db 52 WLKSECRNINRYCDLSAETSDYEHQYAKVKAINGTKCSKWAESGRFPFLETOIGPP 111  
QY 11 -----AEKLE-----16  
Db 112 EVALTTDEKSISVLTAPKWKRNPEPLVSMQIYSNLKYNVSLNRTKSQCVTN 171  
QY 17 ---AFDWL-----21  
Db 172 HTLVLTWLEPNTLYCVHVESFVPGPPRAQPSKQCAKTLKDQSSEASTKGPSVFPPLAPS 231  
QY 22 -----21

Db 232 SKSTSGGTAALGCLVKDYFPEPTVSWNSGALTSGVHTFPAVLQSSGLYSLSSVTVPS 291  
QY 22 -----KAFYDKVAEKLKAFMDKTHTCPPCPAPPELLGGPSVFLPPKPKDPTL 68  
Db 292 SLGTQYICNVNHNKPSNTKVDKKVEPKSCDKTHTCPPCPAPPELLGGPSVFLPPKPKDPTL 351  
QY 69 MISRTPEVTCVVVDVSHEDPEVKFNKYVDGVVHNNAKTPREEQYNSTYRVVSVLTVLHQ 128  
Db 352 MISRTPEVTCVVVDVSHEDPEVKFNKYVDGVVHNNAKTPREEQYNSTYRVVSVLTVLHQ 411  
QY 129 DWLNGKEYCKVSNKALPAPIEKTISKAKGQPREPOVYTLPPSRDELTKNQVSLTCLVK 188  
Db 412 DWLNGKEYCKVSNKALPAPIEKTISKAKGQPREPOVYTLPPSRDELTKNQVSLTCLVK 471  
QY 189 FYPDSIAVWESNGQPENNYKTTPPVLDSDGSFFLYSKLTVDKSRWQQGNVFCSCVMHEA 248  
Db 472 FYPDSIAVWESNGQPENNYKTTPPVLDSDGSFFLYSKLTVDKSRWQQGNVFCSCVMHEA 531  
QY 249 LHHNYTKSLSLSPGK 264  
Db 532 LHHNYTKSLSLSPGK 547

RESULT 11  
US-09-746-359A-53  
; Sequence 53, Application US/09746359A  
; Patent No. US20020042366A1  
; GENERAL INFORMATION:  
; APPLICANT: Thompson, Penny  
; APPLICANT: Foster, Donald C.  
; APPLICANT: Xu, Wenfeng  
; APPLICANT: Madden, Karen L.  
; APPLICANT: Kelly, James D.  
; APPLICANT: Sprecher, Cindy A.  
; APPLICANT: Blumberg, Hal

; APPLICANT: Eagan, Maribeth A.  
; APPLICANT: Jaspers, Stephen R.  
; APPLICANT: Chandrasekher, Yasmin A.  
; APPLICANT: No. US20020042366Alak, Julia E.  
; TITLE OF INVENTION: Method for Treating Inflammation  
; FILE REFERENCE: 99-108  
; CURRENT APPLICATION NUMBER: US/09/746,359A  
; CURRENT FILING DATE: 2001-05-21  
; PRIOR APPLICATION NUMBER: 60/171,969  
; PRIOR FILING DATE: 1999-12-23  
; PRIOR APPLICATION NUMBER: 60/213,341  
; PRIOR FILING DATE: 2000-06-22  
; NUMBER OF SEQ ID NOS: 72  
; SOFTWARE: FastSeq for Windows Version 3.0  
; SEQ ID NO 53  
; LENGTH: 571  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-09-746-359A-53

Query Match 88.0%; Score 1259; DB 10; Length 571;  
Best Local Similarity 48.6%; Pred. No. 5.8e-32;  
Matches 241; Conservative 4; Mismatches 18; Indels 233; Gaps 4;  
QY 2 WLK-----AFYDKV-----10  
Db 76 WLKSECRNINRYCDLSAETSDYEHQYAKVKAINGTKCSKWAESGRFPFLETOIGPP 135  
QY 11 -----AEKLE-----16  
Db 136 EVALTTDEKSISVLTAPKWKRNPEPLVSMQIYSNLKYNVSLNRTKSQCVTN 195  
QY 17 ---AFDWL-----21  
Db 196 HTLVLTWLEPNTLYCVHVESFVPGPPRAQPSKQCAKTLKDQSSEASTKGPSVFPPLAPS 255  
QY 22 -----21

Db 256 SKSTSGGTAALGCLVKDYFPEPTVSWNSGALTSGVHTFPAVLQSSGLYSLSSVTVPS 315  
QY 22 -----KAFYDKVAEKLKAFMDKTHTCPPCPAPPELLGGPSVFLPPKPKDPTL 68  
Db 316 SLGTQYICNVNHNKPSNTKVDKKVEPKSCDKTHTCPPCPAPPELLGGPSVFLPPKPKDPTL 375  
QY 69 MISRTPEVTCVVVDVSHEDPEVKFNKYVDGVVHNNAKTPREEQYNSTYRVVSVLTVLHQ 128  
Db 376 MISRTPEVTCVVVDVSHEDPEVKFNKYVDGVVHNNAKTPREEQYNSTYRVVSVLTVLHQ 435  
QY 129 DWLNGKEYCKVSNKALPAPIEKTISKAKGQPREPOVYTLPPSRDELTKNQVSLTCLVK 188  
Db 436 DWLNGKEYCKVSNKALPAPIEKTISKAKGQPREPOVYTLPPSRDELTKNQVSLTCLVK 495  
QY 189 FYPDSIAVWESNGQPENNYKTTPPVLDSDGSFFLYSKLTVDKSRWQQGNVFCSCVMHEA 248  
Db 496 FYPDSIAVWESNGQPENNYKTTPPVLDSDGSFFLYSKLTVDKSRWQQGNVFCSCVMHEA 555  
QY 249 LHHNYTKSLSLSPGK 264  
Db 556 LHHNYTKSLSLSPGK 571

RESULT 12  
US-09-740-002-27  
; Sequence 27, Application US/09740002  
; Patent No. US20020001798A1  
; GENERAL INFORMATION:  
; APPLICANT: BRAMS, PETER  
; APPLICANT: MORROW, PHILLIP  
; TITLE OF INVENTION: NEUTRALIZING HIGH AFFINITY HUMAN MONOCLONAL ANTIBODIES  
; TITLE OF INVENTION: SPECIFIC TO RSV F-PROTEIN AND METHODS FOR THEIR  
; TITLE OF INVENTION: MANUFACTURE AND THERAPEUTIC USE THEREOF  
; FILE REFERENCE: 037003-0275759  
; CURRENT APPLICATION NUMBER: US/09/740,002

; CURRENT FILING DATE: 2000-12-20  
; PRIOR APPLICATION NUMBER: 09/335,697  
; PRIOR FILING DATE: 1999-06-18  
; PRIOR APPLICATION NUMBER: 08/488,376  
; PRIOR FILING DATE: 1995-06-07  
; NUMBER OF SEQ ID NOS: 27  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 27  
; LENGTH: 475  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-09-740-002-27

Query Match 88.0%; Score 1258; DB 10; Length 475;  
Best Local Similarity 55.3%; Pred. No. 4e-32;  
Matches 236; Conservative 9; Mismatches 12; Indels 170; Gaps 3;  
QY 1 DWLKAFYDKVAEKLKEAFDWL----- 21  
DB 56 NWIR-----QPPGKALEWLARIDWDDDTFYASLTKRLSISKDTSKNQVVLRTNVDP 108  
QY 22 -----KAFYD----- 26  
DB 109 VDTATYFCARASLYDSDFLYEYHAYWGQTVTVSSASTKGPSVFPLAPSKSTSGGTA 168  
QY 27 ----- 26  
DB 169 ALGCLVKDYFPEPTVSNNGALTSGVHTFPAVLQSSGLYSLSSVTVTPSSSLGTQTYIC 228  
QY 27 -----KVAEKLKEAFMDKTHCTCPCPAPPELLGGPSVFLFPPKPKDTLMISRTPEVT 77  
DB 229 NVNHRKPSNTKVDKAEKPSKCDKTHCTCPCPAPPELLGGPSVFLFPPKPKDTLMISRTPEVT 288  
QY 78 CVVDVDSHEDPEVKFNWYVDGVEVHNAKTKPREEQYNSTYRVVSVLTVLHQDWLNGKEYK 137  
DB 289 CVVDVDSHEDPEVKFNWYVDGVEVHNAKTKPREEQYNSTYRVVSVLTVLHQDWLNGKEYK 348  
QY 138 KVSNNKALPAPIEKTISKAKGQPREPOVYVTLPPSRDELTKNQVSLTCLVKGFYPSDIAVE 197  
DB 349 KVSNNKALPAPIEKTISKAKGQPREPOVYVTLPPSRDELTKNQVSLTCLVKGFYPSDIAVE 408  
QY 198 WESNQGPNYKTTTPVLDSDGSFFLYSKLTVDKSRWQGNVFSCVMHEALHNHYTQKS 257  
DB 409 WESNQGPNYKTTTPVLDSDGSFFLYSKLTVDKSRWQGNVFSCVMHEALHNHYTQKS 468  
QY 258 LSLSPGK 264  
DB 469 LSLSPGK 475

RESULT 13  
US-10-091-236-17  
; Sequence 17, Application US/10091236  
; Patent No. US20020168360A1  
; GENERAL INFORMATION:  
; APPLICANT: DINGIVAN, CHRISTINE A.  
; TITLE OF INVENTION: METHODS OF PREVENTING OR TREATING INFLAMMATORY OR AUTOIMMUNE  
; TITLE OF INVENTION: DISORDERS BY ADMINISTERING INTEGRIN ALPHA-V-BETA-3 ANTAGONISTS  
; TITLE OF INVENTION: COMBINATION WITH OTHER PROPHYLACTIC OR THERAPEUTIC AGENTS  
; FILE REFERENCE: 10271-053-999  
; CURRENT APPLICATION NUMBER: US/10/091,236  
; CURRENT FILING DATE: 2002-03-04  
; PRIOR APPLICATION NUMBER: US 60/273,098  
; PRIOR FILING DATE: 2001-03-02  
; PRIOR APPLICATION NUMBER: US 60/316,321  
; PRIOR FILING DATE: 2001-08-31  
; NUMBER OF SEQ ID NOS: 17  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 17  
; LENGTH: 347  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-10-091-236-17

Query Match 87.8%; Score 1256; DB 9; Length 347;  
Best Local Similarity 78.5%; Pred. No. 2.3e-32;  
Matches 238; Conservative 3; Mismatches 11; Indels 51; Gaps 3;  
QY 2 WLKAFYDKVAEKLKEAF-----DWL 21  
DB 56 WKKQ-KDKVAELEENSEFFRAFSFKNRVYLDTVYSGSLTIYNLTSSDEDEYEMESPNTIDTM 114  
QY 22 KAFYDKVAEKLKEAFMDKTHCTCPCPAPPELLGGPSVFLFPPKPKDTLMISRTPEVTCVV 81  
DB 115 KFF-----LYVDKTHCTCPCPAPPELLGGPSVFLFPPKPKDTLMISRTPEVTCVV 164  
QY 82 DVSHEDEPEVKFNWYVDGVEVHNAKTKPREEQYNSTYRVVSVLTVLHQDWLNGKEYKCKYS 141  
DB 165 DVSHEDEPEVKFNWYVDGVEVHNAKTKPREEQYNSTYRVVSVLTVLHQDWLNGKEYKCKYS 224  
QY 142 NKALPAPIEKTISKAKGQPREPOVYVTLPPSRDELTKNQVSLTCLVKGFYPSDIAVEWESN 201  
DB 225 NKALPAPIEKTISKAKGQPREPOVYVTLPPSRDELTKNQVSLTCLVKGFYPSDIAVEWESN 284  
QY 202 GOPENNYKTTTPVLDSDGSFFLYSKLTVDKSRWQGNVFSCVMHEALHNHYTQKSLSLS 261  
DB 285 GOPENNYKTTTPVLDSDGSFFLYSKLTVDKSRWQGNVFSCVMHEALHNHYTQKSLSLS 344  
QY 262 PGK 264  
DB 345 PGK 347

RESULT 14  
US-10-091-313-7  
; Sequence 7, Application US/10091313  
; Publication No. US20030044406A1  
; GENERAL INFORMATION:  
; APPLICANT: DINGIVAN, CHRISTINE  
; TITLE OF INVENTION: METHODS OF PREVENTING OR TREATING INFLAMMATORY OR AUTOIMMUNE  
; TITLE OF INVENTION: DISORDERS BY ADMINISTERING CD2 ANTAGONISTS IN COMBINATION WITH  
; TITLE OF INVENTION: PROPHYLACTIC OR THERAPEUTIC AGENTS  
; FILE REFERENCE: 10271-063  
; CURRENT APPLICATION NUMBER: US/10/091,313  
; CURRENT FILING DATE: 2002-03-04  
; PRIOR APPLICATION NUMBER: US 60/273,098  
; PRIOR FILING DATE: 2001-03-02  
; PRIOR APPLICATION NUMBER: US 60/346,918  
; PRIOR FILING DATE: 2001-10-19  
; NUMBER OF SEQ ID NOS: 13  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 7  
; LENGTH: 347  
; TYPE: PRT  
; ORGANISM: Homo Sapiens  
US-10-091-313-7  
Query Match 87.8%; Score 1256; DB 9; Length 347;  
Best Local Similarity 78.5%; Pred. No. 2.3e-32;  
Matches 238; Conservative 3; Mismatches 11; Indels 51; Gaps 3;  
QY 2 WLKAFYDKVAEKLKEAF-----DWL 21  
DB 56 WKKQ-KDKVAELEENSEFFRAFSFKNRVYLDTVYSGSLTIYNLTSSDEDEYEMESPNTIDTM 114  
QY 22 KAFYDKVAEKLKEAFMDKTHCTCPCPAPPELLGGPSVFLFPPKPKDTLMISRTPEVTCVV 81  
DB 115 KFF-----LYVDKTHCTCPCPAPPELLGGPSVFLFPPKPKDTLMISRTPEVTCVV 164  
QY 82 DVSHEDEPEVKFNWYVDGVEVHNAKTKPREEQYNSTYRVVSVLTVLHQDWLNGKEYKCKYS 141  
DB 165 DVSHEDEPEVKFNWYVDGVEVHNAKTKPREEQYNSTYRVVSVLTVLHQDWLNGKEYKCKYS 224  
QY 142 NKALPAPIEKTISKAKGQPREPOVYVTLPPSRDELTKNQVSLTCLVKGFYPSDIAVEWESN 201  
DB 225 NKALPAPIEKTISKAKGQPREPOVYVTLPPSRDELTKNQVSLTCLVKGFYPSDIAVEWESN 284

[illegible]

RESULT 15  
 US-10-091-268-7  
 ; Sequence 7, Application US/10091268  
 ; Publication No. US20030068320A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: DINGIVAN, CHRISTINE A  
 ; TITLE OF INVENTION: METHODS OF ADMINISTERING/DOING CD2 ANTAGONISTS FOR THE PREVENTION  
 ; OF DISEASES OF THE IMMUNE SYSTEM  
 ; FILE REFERENCE: 10271-054-999  
 ; CURRENT APPLICATION NUMBER: US/10/091,268  
 ; CURRENT FILING DATE: 2002-03-04  
 ; PRIOR APPLICATION NUMBER: US 60/273,098  
 ; PRIOR FILING DATE: 2001-03-02  
 ; PRIOR APPLICATION NUMBER: US 60/346,918  
 ; PRIOR FILING DATE: 2001-10-19  
 ; NUMBER OF SEQ ID NOS: 7  
 ; SOFTWARE: PatentIn version 3.1  
 ; SEQ ID NO 7  
 ; .LENGTH: 347  
 ; TYPE: PRT  
 ; ORGANISM: Homo sapiens  
 ; US-10-091-268-7

```

Query Match      87.8%; Score 1256; DB 9; Length 347;
Best Local Similarity 78.5%; Pred. No. 2.3e-32;
Matches 238; Conservative 3; Mismatches 11; Indels 51; Gaps 3;

QY      2 WLKAFYDKVAEKLKEAF-----DWL 21
      ||| ||||| |
Db      56 WKQK-KDKVAELENSEPFAFSSFKNRVYLTIVTSGSLTYINLTSSDEDEYEMESPNIITDTM 114
      ||| ||||| |
QY      22 KAFYDKVAEKLKEAFMDKTHTCPCPCAPPELLGGPSVFLFPPKPKDTHLMISRTPEVTCVV 81
      ||| ||||| |
Db      115 KFF-----LYVDKTHTCPCPCAPPELLGGPSVFLFPPKPKDTHLMISRTPEVTCVV 164
      ||| ||||| |
QY      82 DVSHEDPEVFNNYVDGVEVHNNAKTKPREQYNSTYRWVSVLTVLHQDWLNGKEYCKVKS 141
      ||| ||||| |
Db      165 DVSHEDPEVFNNYVDGVEVHNNAKTKPREQYNSTYRWVSVLTVLHQDWLNGKEYCKVKS 224
      ||| ||||| |
QY      142 NKALPAPIEKTISKAKGQPREPQVYTLPPSRDELTKNQVSLTCLVKGFYPSDIAVEWESN 201
      ||| ||||| |
Db      225 NKALPAPIEKTISKAKGQPREPQVYTLPPSRDELTKNQVSLTCLVKGFYPSDIAVEWESN 284
      ||| ||||| |
QY      202 GQPNENYKTTTPPVLDSGSEFFLYSKLTVDKSRWQQGNVFSCSVMHEALHNHYTQKSLSLS 261
      ||| ||||| |
Db      285 GQPNENYKTTTPPVLDSGSEFFLYSKLTVDKSRWQQGNVFSCSVMHEALHNHYTQKSLSLS 344
      ||| ||||| |
QY      262 PGK 264
      |||
Db      345 PGK 347

```

Search completed: April 21, 2003, 10:51:19  
Job time : 14.1647 secs



GenCore version 5.1.4.p5.4578  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: April 21, 2003, 10:42:26 ; Search time 19.4118 Seconds  
(without alignments)  
1307.428 Million cell updates/sec

Title: 7LINK7LINK2

Perfect score: 1430

Sequence: 1 DWLKAFYDKVAEKLKEAFDW.....MHEALHNYTKLSLSPGK 264

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.0

Searched: 283224 seqs, 96134422 residues

Total number of hits satisfying chosen parameters: 283224

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

PIR\_73:\*

1: pir1:\*

2: pir2:\*

3: pir3:\*

4: pir4:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1243	86.9	330	1 GHU	Ig gamma-1 chain C
2	1242	86.9	374	2 S69339	Ig heavy chain V r
3	1233	86.2	255	3 S1866	Ig gamma-1 chain C
4	1190	83.2	234	2 P70207	Ig gamma chain C r
5	1165	81.5	377	2 A23511	Ig gamma-3 chain C
6	1163	81.3	377	2 A60764	Ig gamma-3 chain C
7	1148	80.3	326	1 G2HU	Ig gamma-2 chain C
8	1144	80.0	327	1 G4HU	Ig gamma-4 chain C
9	1143	79.9	289	1 G3HUW1	Ig gamma-3 heavy c
10	923	64.5	323	1 GHRB	Ig gamma chain C r
11	917	64.1	328	2 I47160	Ig gamma 2b chain
12	917	64.1	328	2 I47159	Ig gamma 2a chain
13	911	63.7	277	2 I47162	Ig gamma 4 chain C
14	899	62.9	329	1 G2GP	Ig gamma-2 chain C
15	889	62.2	328	2 I47158	Ig gamma 1 chain c
16	885	61.9	328	2 I47161	Ig gamma 3 chain c
17	874	61.1	470	2 S22080	Ig heavy chain pre
18	874	61.1	472	2 S31459	Ig gamma-1 chain -
19	854	59.7	308	2 C30554	Ig heavy chain C r
20	853	59.7	329	1 G3MSC	Ig gamma-3 chain C
21	850	59.4	444	2 PC4436	Ig gamma-2b chain
22	849	59.4	333	2 PS0018	Ig gamma-3 chain C
23	842	58.9	398	1 G3MSM	Ig gamma-2a chain
24	831	58.1	469	2 S37483	Ig gamma-1 chain C
25	826	57.8	326	2 PS0017	Ig gamma-1 chain C
26	823	57.6	324	1 G1MS	Ig gamma-1 chain C
27	823	57.6	446	2 S40295	Ig gamma-2a chain
28	818	57.2	329	2 S00847	Ig gamma-2c chain
29	818	57.2	393	1 G1MSM	Ig gamma-1 chain C

30	815	57.0	330	1 G2MSA	Ig gamma-2a chain
31	810	56.6	399	1 G2MSAM	Ig gamma-2a chain
32	809	56.6	335	1 G2MSAB	Ig gamma-2a chain
33	807	56.4	474	1 G2MS11	Ig gamma-2b chain
34	795	55.6	322	2 PS0019	Ig gamma-2a chain
35	785	54.9	475	2 S01321	Ig gamma-2b chain
36	784	54.8	405	1 G2MSBM	Ig gamma-2b chain
37	774	54.1	327	2 S06611	Ig gamma-2 chain C
38	707	49.4	180	2 I46732	Ig gamma heavy cha
39	605	42.3	218	2 A36040	Ig heavy chain V-I
40	601	42.0	249	2 S69340	Ig heavy chain VHI
41	571	39.9	152	2 S14236	Ig gamma-1 chain C
42	485	33.9	26926	1 I38344	titin, cardiac mus
43	444	31.0	572	2 B46529	Ig y heavy chain (
44	435	30.4	548	2 S38864	Ig epsilon chain C
45	413	28.9	448	2 S03186	Ig heavy chain C r

#### ALIGNMENTS

##### RESULT 1

GHU

Ig gamma-1 chain C region - human

C:Species: Homo sapiens (man)

C:Date: 31-Jan-1981 #sequence\_revision 18-Aug-1982 #text\_change 16-Jul-1999

C:Accession: A93433; S36861; S33887; B90563; A90564; B91668; A91723; A02146

R:Ellison, J.W.; Berson, B.J.; Hood, L.E.

Nucleic Acids Res. 10, 4071-4079, 1982

A:Title: The nucleotide sequence of a human immunoglobulin C-gamma1 gene.

A:Reference number: A93433; MUID:82274238; PMID:6287432

A:Accession: A93433

A:Molecule type: DNA

A:Residues: 1-330 <ELL>

A:Cross-references: EMBL:217370

A:Note: this sequence has the G1m(17) allotypic marker, 97-Lys, and the G1m(1) marker

A:Note: Lys-330 is removed after translation

R:Harris, L.J.

submitted to the EMBL Data Library, October 1992

A:Reference number: S33904

A:Accession: S36861

A:Molecule type: DNA

A:Residues: 2-330 <HAR>

A:Cross-references: EMBL:217370

R:Takahashi, N.; Ueda, S.; Obata, M.; Nikaido, T.; Nakai, S.; Honjo, T.

Cell 29, 671-679, 1982

A:Title: Structure of human immunoglobulin gamma genes: implications for evolution of

A:Reference number: S33887; MUID:83001943; PMID:6811139

A:Accession: S33887

A:Molecule type: DNA

A:Residues: 88-113;235-330 <TAK>

A:Cross-references: EMBL:217370

R:Cunningham, B.A.; Rutishauser, U.; Gall, W.E.; Gottlieb, P.D.; Waxdal, M.J.; Edelman

Biochemistry 9, 3161-3170, 1970

A:Title: The covalent structure of a human gamma-gamma-immunoglobulin. VII. Amino acid seq

A:Reference number: A90563; MUID:71064024; PMID:5489771

A:Contents: myeloma protein Eu

A:Accession: B90563

A:Molecule type: protein

A:Residues: 1-96,'R',98-135 <CUN>

A:Note: this sequence has the G1m(3) marker, 97-Arg

R:Rutishauser, U.; Cunningham, B.A.; Bennett, C.; Konigsberg, W.H.; Edelman, G.M.

Biochemistry 9, 3171-3181, 1970

A:Title: The covalent structure of a human gamma-gamma-immunoglobulin. VIII. Amino acid se

A:Reference number: A90564; MUID:71064025; PMID:5530842

A:Contents: Eu

A:Accession: A90564

A:Molecule type: protein

A:Residues: 136-154,'Q',156-165,'Q',167-176,'Q',178-194,'N',196-197,'D',199-238,'E',2

A:Note: this sequence has the G1m(non-1) markers, 239-Glu and 241-Met

R:Ponstingl, H.; Hilschmann, N.

Hoppe-Seyler's Z. Physiol. Chem. 357, 1571-1604, 1976

A:Title: Die Primaerstruktur eines monoklonalen IgG1-Immunglobulins (Myelomprotein Ni

## Igen Primaerstruktur.

A:Reference number: A91668; MUID:77070269; PMID:826475  
A:Contents: myeloma protein Nie  
A:Accession: B91668  
A:Molecule type: protein  
A:Residues: 1-34, 'Q', '36-96', 'K', '98-115', 'Q', '117-197', 'D', '199-238', 'D', '240', 'L', '242-268', 'E', '27  
A:Note: this sequence has the Gln(17) and Gln(1) markers  
R:Schmidt, W.E.; Jung, H.D.; Palm, W.; Hilschmann, N.  
Hoppe-Seyler's Z. Physiol. Chem. 364, 713-747, 1983  
A:Title: Die Primaerstruktur des kristallisierten monoklonalen Immunglobulins IgG1 KOL  
A:Reference number: A91723; MUID:83289131; PMID:6884994  
A:Contents: myeloma protein KOL; disulfide bonds  
A:Accession: A91723  
A:Molecule type: protein  
A:Residues: 1-96, 'R', '98-197', 'D', '199-238', 'E', '240', 'M', '242-266', 'D', '268-271', 'D', '273-330' <SCH  
A:Note: this sequence has the Gln(3) and Gln(non-1) markers  
B:Gall, W.E.; Edelman, G.M.  
Biochemistry 9, 3188-3196, 1970  
A:Title: The covalent structure of a human gammaG-immunoglobulin. X. Intrachain disulfide  
A:Reference number: A90565; MUID:71064027; PMID:4923144  
A:Contents: annotation; disulfide bonds  
R:Dreker, L.; Schwarz, J.; Reichel, W.; Hilschmann, N.  
Hoppe-Seyler's Z. Physiol. Chem. 357, 1515-1540, 1976  
A:Title: Rule of antibody structure. The primary structure of monoclonal IgG1 immunoglob  
enamide cleavage products, and the disulfide bridges.  
A:Reference number: A91667; MUID:77070267; PMID:1002129  
A:Contents: annotation; disulfide bonds  
C:Genetics:  
A:Gene: GDB:IGHG1  
A:Cross-references: GDB:120085; OMIM:147100  
A:Map position: 14q32.33-14q32.33  
A:Introns: 99/1; 114/1; 224/1  
C:Complex: An immunoglobulin heterotetramer subunit consists of two identical light (kap  
hain disulfide bonds. In some cases, such as IgA and IgM, the subunits associate into la  
C:Superfamily: immunoglobulin C region; immunoglobulin homology  
C:Keywords: duplication; glycoprotein; heterotetramer; immunoglobulin  
F:20-85/Domain: immunoglobulin homology <IMI>  
F:137-206/Domain: immunoglobulin homology <IM2>  
F:243-310/Domain: immunoglobulin homology <IM3>  
F:27-83,144-204,250-308/Disulfide bonds: #status experimental  
F:103/Disulfide bonds: Interchain (to light chain) #status experimental  
F:109,112/Disulfide bonds: interchain (to heavy chain) #status experimental  
F:180/Binding site: carbohydrate (Asn) (covalent) #status experimental  
Query Match 86.9%; Score 1243; DB 1; Length 330;  
Best Local Similarity 96.6%; Pred. No. 9.6e-34;  
Matches 230; Conservative 3; Mismatches 5; Indels 0; Gaps 0;  
QY 27 KVAEKLKAFMDKTHTCPCPAPELLGGPSVFLPPKPKDTLMISRTPEVTCVVVDVSH 86  
Db 93 KVDKKVEPKSCDKTHTCPCPAPELLGGPSVFLPPKPKDTLMISRTPEVTCVVVDVSH 152  
QY 87 DPEVKFNKYVDGVEVHNKAKTPREEQYNSTYRVYSVLTVLDHQLNGKEYKCKVSNKALP 146  
Db 153 DPEVKFNKYVDGVEVHNKAKTPREEQYNSTYRVYSVLTVLDHQLNGKEYKCKVSNKALP 212  
QY 147 APIETISKAKGPQREPOVYTLPPSRDELTKNQVSLTCLVKGFYPSDIAVWESNGQPEN 206  
Db 213 APIETISKAKGPQREPOVYTLPPSRDELTKNQVSLTCLVKGFYPSDIAVWESNGQPEN 272  
QY 207 NYKTPPVLDSDGSFFLYSKLTVDKSRWQQGNVSCSVMHHEALHNHYTQKSLSLSPGK 264  
Db 273 NYKTPPVLDSDGSFFLYSKLTVDKSRWQQGNVSCSVMHHEALHNHYTQKSLSLSPGK 330  
RESULT 2  
S69339  
Ig heavy chain V region precursor - human  
C:Species: Homo sapiens (man)  
C:Date: 19-Mar-1997 #sequence\_revision 19-Mar-1997 #text\_change 01-Dec-2000  
A:Accession: S69339; S72664  
R:Khamilich, A.A.; Aucouturier, P.; Preud'homme, J.L.; Cogné, M.  
Eur. J. Biochem. 229, 54-60, 1995

A:Title: Structure of abnormal heavy chains in human heavy-chain-deposition disease.  
A:Reference number: S69339; MUID:95262687; PMID:7744049

A:Accession: S69339  
A>Status: preliminary  
A:Molecule type: mRNA  
A:Residues: 1-374 <KHA>  
A:Cross-references: EMBL:X81695  
R:Khamilich, A.A.  
submitted to the EMBL Data Library, September 1994  
A:Reference number: S72664  
A:Accession: S72664  
A>Status: preliminary  
A:Molecule type: mRNA  
A:Residues: 1-140, 'C', '142-374' <KH2>  
A:Cross-references: EMBL:X81695  
C:Superfamily: immunoglobulin C region; immunoglobulin homology

Query Match 86.9%; Score 1242; DB 2; Length 374;  
Best Local Similarity 71.3%; Pred. No. 1.6e-33;  
Matches 233; Conservative 10; Mismatches 11; Indels 73; Gaps 4;

QY 2 WLKAFYDKVAEKLKAEFDWLKAFY---DK----- 27  
Db 57 WIR-----QPPQALEWLALIFWDDDKRYSPSLRTRLTITKTSKNQVLTWTVNDPA 109  
QY 28 -----VAEKLKAEAFMDKTHTCPCPAPELLGGPSV 57  
Db 110 DTATYYCGYSVEGYGQGRFHSWGQGLTVTVSSSEPKSC--DKTHTCPCPAPELLGGPSV 167  
QY 58 FLPPPKPKDTLMISRTPEVTCVVVDVSHEDPEVKFNKYVDGVEVHNKAKTPREEQYNSTY 117  
Db 168 FLPPPKPKDTLMISRTPEVTCVVVDVSHEDPEVKFNKYVDGVEVHNKAKTPREEQYNSTY 227  
QY 118 RVSVLTCLVKGFYPSDIAVWESNGQPENNYKTPPVLDSDGSFFLYSKLTVDKSRWQQG 177  
Db 228 RVSVLTCLVKGFYPSDIAVWESNGQPENNYKTPPVLDSDGSFFLYSKLTVDKSRWQQG 287  
QY 178 NVSLTCLVKGFYPSDIAVWESNGQPENNYKTPPVLDSDGSFFLYSKLTVDKSRWQQG 237  
Db 288 NVSLTCLVKGFYPSDIAVWESNGQPENNYKTPPVLDSDGSFFLYSKLTVDKSRWQQG 347  
QY 238 NVFSCSYMHHEALHNHYTQKSLSLSPGK 264  
Db 348 NVFSCSYMHHEALHNHYTQKSLSLSPGK 374

## RESULT 3

S31866  
Ig gamma-1 chain C region - synthetic

C:Species: synthetic  
A:Note: Homo sapiens (man) gene engineered and expressed in *Escherichia coli*  
C:Date: 06-Jan-1995 #sequence\_revision 17-Mar-1997 #text\_change 19-May-2000  
C:Accession: S31866

R:Filipula D.

submitted to the EMBL Data Library, February 1993

A:Description: Screening method for protein-protein interactions of cloned gene produc  
A:Reference number: S31866

A:Accession: S31866

A:Molecule type: mRNA

A:Residues: 1-255 <FIL>

A:Cross-references: EMBL:X70421; NID:g33068; PIDN:CAA49866.1; PID:g33069

C:Keywords: Immunoglobulin

F:1-22/Region: *Escherichia coli* outer membrane protein A precursor

F:23-255/Region: human Ig gamma-1 chain C region

Query Match 86.2%; Score 1233; DB 4; Length 255;  
Best Local Similarity 100.0%; Pred. No. 8.7e-34;  
Matches 227; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 38 DKTHTCPCPAPELLGGPSVFLPPKPKDTLMISRTPEVTCVVVDVSHEDPEVKFNKYVD 97

Db 29 DKTHTCPCPAPELLGGPSVFLPPKPKDTLMISRTPEVTCVVVDVSHEDPEVKFNKYVD 88

QY 98 GVEVHNAKTPREQYNSTYRVVSVTLVHLDQWLNKCKYKCKVSNKALPAPIEKTISKAK 157  
Db 89 GVEVHNAKTPREQYNSTYRVVSVTLVHLDQWLNKCKYKCKVSNKALPAPIEKTISKAK 148  
QY 158 GQPREQVYTLPPSRDELTKNQVSLTCLVKGFYPSDIAVWESNGQPENNYKTTPPVLD 217  
Db 149 GQPREQVYTLPPSRDELTKNQVSLTCLVKGFYPSDIAVWESNGQPENNYKTTPPVLD 208  
QY 218 DGSFFLYSKLTVDKSRWQOGNPFSCVMHEALHNHYTQKSLSLSPGK 264  
Db 209 DGSFFLYSKLTVDKSRWQOGNPFSCVMHEALHNHYTQKSLSLSPGK 255

RESULT 4  
PT0207  
Ig gamma chain C region - chimpanzee  
C:Species: Pan troglodytes (chimpanzee)  
C:Date: 23-Nov-1991 #sequence\_revision 23-Nov-1991 #text\_change 16-Jul-1999  
C:Accession: PT0207  
R:Ehrlich, P.H.; Moustafa, Z.A.; Oestberg, L.  
Mol. Immunol. 28, 319-322, 1991  
A:Title: Nucleotide sequence of chimpanzee Fc and hinge regions.  
A:Reference number: PT0207; MUID:91287716; PMID:2062315  
A:Accession: PT0207  
A:Molecule type: mRNA  
A:Residues: 1-234 <EHR>  
C:Superfamily: immunoglobulin C region; immunoglobulin homology  
C:Keywords: immunoglobulin  
F:48-117/Domain: immunoglobulin homology <IMM>

Query Match 83.2%; Score 1190; DB 2; Length 234;  
Best Local Similarity 95.2%; Pred. No. 1.8e-32;  
Matches 220; Conservative 4; Mismatches 7; Indels 0; Gaps 0;

QY 27 KVAEKLKEAFMDKTHTCPCPAPPELLGGPSVFLFPPPKDPTLMISRTPEVTCVVDVSH 86  
Db 4 KVDKKEPKSCDTHTCPCPAPPELLGGPSVFLFPPPKDPTLMISRTPEVTCVVDVSH 63

QY 87 DPEVKFNWYVDGVEVHNAKTPREQYNSTYRVVSVTLVHLDQWLNKCKYKCKVSNKALP 146  
Db 64 DPEVKFNWYVDGVEVHNAKTPREQYNSTYRVVSVTLVHLDQWLNKCKYKCKVSNKALP 123

QY 147 APIEKTISKAKGQPREPQVYTLPPSRDELTKNQVSLTCLVKGFYPSDIAVWESNGQPEN 206  
Db 124 APIEKTISKAKGQPREPQVYTLPPSRDELTKNQVSLTCLVKGFYPSDIAVWESNGQPEN 183

QY 207 NYKTTPPVLDSDGSFFLYSKLTVDKSRWQOGNPFSCVMHEALHNHYTQKS 257  
Db 184 NYKTTPPVLDSDGSFFLYSKLTVDKSRWQOGNPFSCVMHEALHNHYTQKS 234

RESULT 5  
A23511  
Ig gamma-3 chain C region (allotype G3m(b)) - human  
C:Species: Homo sapiens (man)  
C:Date: 28-Dec-1987 #sequence\_revision 28-Dec-1987 #text\_change 23-Jul-1999  
C:Accession: A23511  
R:Huck, S.; Fort, P.; Crawford, D.H.; Lefranc, M.P.; Lefranc, G.  
Nucleic Acids Res. 14, 1779-1789, 1986  
A:Title: Sequence of a human immunoglobulin gamma 3 heavy chain constant region gene: c  
A:Reference number: A23511; MUID:86148507; PMID:3081877  
A:Accession: A23511  
A:Molecule type: DNA  
A:Residues: 1-377 <HUC>  
A:Cross-references: GB:X03604; GB:M12958; MID:g33070; PIDN:CAA27268.1; PID:g577056  
C:Genetics:  
A:Gene: GDB:IGHG3  
A:Cross-references: GDB:119339; OMIM:147120  
A:Map position: 14q32.33-14q32.33  
A:Introns: 98/3; 115/3; 130/3; 145/3; 160/3; 270/3  
C:Superfamily: immunoglobulin C region; immunoglobulin homology  
C:Keywords: immunoglobulin  
F:20-85/Domain: immunoglobulin homology <IMM>

Query Match 81.5%; Score 1165; DB 2; Length 377;  
Best Local Similarity 76.8%; Pred. No. 6.1e-31;  
Matches 218; Conservative 8; Mismatches 12; Indels 46; Gaps 2;

QY 26 DKVAEKLKEAFMDKTHT----- 42  
Db 95 DKRVE-LKTPLDGDTHTCPCPAPPELLGGPSVFLFPPPKDPTLMISRTPEVTCVVDVSH 153

QY 43 --CPCPAPPELLGGPSVFLFPPPKDPTLMISRTPEVTCVVDVSHEDPEVKFNWYVDGVE 100  
Db 154 PCPCPAPPELLGGPSVFLFPPPKDPTLMISRTPEVTCVVDVSHEDPEVKFNWYVDGVE 213

QY 101 VHNAKTPREQYNSTYRVVSVTLVHLDQWLNKCKYKCKVSNKALPAPIEKTISKAKGQP 160  
Db 214 VHNAKTPREQYNSTYRVVSVTLVHLDQWLNKCKYKCKVSNKALPAPIEKTISKAKGQP 273

QY 161 REPQVYTLPPSRDELTKNQVSLTCLVKGFYPSDIAVWESNGQPENNYKTTPPVLDSDGS 220  
Db 274 REPQVYTLPPSRDELTKNQVSLTCLVKGFYPSDIAVWESNGQPENNYKTTPPVLDSDGS 333

QY 221 FFYLYSKLTVDKSRWQOGNPFSCVMHEALHNHYTQKSLSLSPGK 264  
Db 334 FFYLYSKLTVDKSRWQOGNPFSCVMHEALHNHYTQKSLSLSPGK 377

RESULT 6  
A60764  
Ig gamma-3 chain C region, form LAT - human  
C:Species: Homo sapiens (man)  
C:Date: 14-May-1993 #sequence\_revision 14-May-1993 #text\_change 16-Jul-1999  
C:Accession: A60764  
R:Huck, S.; Lefranc, G.; Lefranc, M.P.  
Immunogenetics 30, 250-257, 1989  
A:Title: A human immunoglobulin IGHG3 allele (Gmb0, bl, c3; c5, u) with an IGHG4 con  
A:Reference number: A60764; MUID:90007613; PMID:2571587  
A:Accession: A60764  
A:Molecule type: DNA  
A:Residues: 1-377 <HUC>  
C:Superfamily: immunoglobulin C region; immunoglobulin homology  
C:Keywords: immunoglobulin  
F:20-85/Domain: immunoglobulin homology <IMM>

Query Match 81.3%; Score 1163; DB 2; Length 377;  
Best Local Similarity 76.8%; Pred. No. 7.1e-31;  
Matches 218; Conservative 8; Mismatches 12; Indels 46; Gaps 2;

QY 26 DKVAEKLKEAFMDKTHT----- 42  
Db 95 DKRVE-LKTPLDGDTHTCPCPAPPELLGGPSVFLFPPPKDPTLMISRTPEVTCVVDVSH 153

QY 43 --CPCPAPPELLGGPSVFLFPPPKDPTLMISRTPEVTCVVDVSHEDPEVKFNWYVDGVE 100  
Db 154 PCPCPAPPELLGGPSVFLFPPPKDPTLMISRTPEVTCVVDVSHEDPEVKFNWYVDGVE 213

QY 101 VHNAKTPREQYNSTYRVVSVTLVHLDQWLNKCKYKCKVSNKALPAPIEKTISKAKGQP 160  
Db 214 VHNAKTPREQYNSTYRVVSVTLVHLDQWLNKCKYKCKVSNKALPAPIEKTISKAKGQP 273

QY 161 REPQVYTLPPSRDELTKNQVSLTCLVKGFYPSDIAVWESNGQPENNYKTTPPVLDSDGS 220  
Db 274 REPQVYTLPPSRDELTKNQVSLTCLVKGFYPSDIAVWESNGQPENNYKTTPPVLDSDGS 333

QY 221 FFYLYSKLTVDKSRWQOGNPFSCVMHEALHNHYTQKSLSLSPGK 264  
Db 334 FFYLYSKLTVDKSRWQOGNPFSCVMHEALHNHYTQKSLSLSPGK 377

RESULT 7  
G2HU  
Ig gamma-2 chain C region - human  
C:Species: Homo sapiens (man)

C>Date: 30-Apr-1981 #sequence\_revision 13-Jun-1983 #text\_change 21-Jul-2000  
C:Accession: A93906; A92809; A90752; A93132; A02148  
R:Ellison, J.; Hood, L.  
Proc. Natl. Acad. Sci. U.S.A. 79, 1984-1988, 1982  
A:Title: Linkage and sequence homology of two human immunoglobulin gamma heavy chain con  
A:Reference number: A93906; MUID:82197621; PMID:6804948  
A:Accession: A93906  
A:Molecule type: DNA  
A:Residues: 1-326 <ELL>  
A:CROSS-references: GB:U00554; GB:J00230; NID:g32759; PIDN:CAB58438.1; PID:g6066056  
A:Note: Lys-326 is probably removed posttranslationally  
R:Wang, A.C.; Tung, E.; Eudenberg, H.H.  
J. Immunol. 125, 1048-1054, 1980  
A:Title: The primary structure of a human IgG2 heavy chain: genetic, evolutionary, and s  
A:Reference number: A92809; MUID:81007873; PMID:6774012  
A:Contents: myeloma protein T11  
A:Accession: A92809  
A:Molecule type: protein  
A:Residues: 1-19, 'Q', 21-57, 'Z', 59, 'A', 61-193, 'D', 195-325 <WAN>  
A:Note: Trp-156 is at or near the complement-binding site  
R:Connell, G.E.; Parr, D.M.; Hofmann, T.  
Can. J. Biochem. 57, 758-767, 1979  
A:Title: The amino acid sequences of the three heavy chain constant region domains of a  
A:Reference number: A90752; MUID:80001357; PMID:113060  
A:Contents: myeloma protein Zie  
A:Accession: A90752  
A:Molecule type: protein  
A:Residues: 1-24, 'E', 26-57, 'EV', 60-85; 132-171, 'zzz', 175, 'B', 177-193, 'D', 195-196, 'Q', 198-  
A:Note: This sequence has since been revised  
R:Hofmann, T.; Parr, D.M.  
Mol. Immunol. 16, 923-925, 1979  
A:Title: A note on the amino acid sequence of residues 381-391 of human immunoglobulin g  
A:Reference number: A93132; MUID:80114419; PMID:118920  
A:Contents: Zie  
A:Accession: A93132  
A:Molecule type: protein  
A:Residues: 238-275 <HOF>  
R:Hofmann, T.; Parr, D.M.  
submitted to the Atlas, March 1980  
A:Reference number: A94591  
A:Contents: annotation: Zie, revisions to residues 25, 59, 60, and 264-268  
A:Note: the revised sequence differs from that shown in having 60-Ala and in the amidati  
ned  
R:Milstein, C.; Frangione, B.  
Biochem. J. 121, 217-225, 1971  
A:Title: Disulfide bridges of the heavy chain of human immunoglobulin G2.  
A:Reference number: A90253; MUID:72033500; PMID:4940472  
A:Contents: annotation: myeloma protein Sa, disulfide bonds  
R:Frangione, B.; Milstein, C.; Pink, J.R.L.  
Nature 221, 145-148, 1969  
A:Title: Structural studies of immunoglobulin G.  
A:Reference number: A93157; MUID:69064124; PMID:5782707  
A:Contents: annotation: Sa, disulfide bonds  
C:Genetics:  
A:Gene: GDB:IGHG2  
A:CROSS-references: GDB:119338; OMIM:147110  
A:Map position: 14q32.33-14q32.33  
C:Complex: An immunoglobulin heterotetramer subunit consists of two identical light (kap  
hain disulfide bonds. In some cases, such as IgA and IgM, the subunits associate into la  
C:Superfamily: immunoglobulin C region; immunoglobulin homology  
C:Keywords: duplication; glycoprotein; heterotetramer; immunoglobulin  
F:20-85/Domain: immunoglobulin homology <IM1>  
F:133-202/Domain: immunoglobulin homology <IM2>  
F:239-306/Domain: immunoglobulin homology <IM3>  
F:14/Disulfide bonds: interchain (to light chain) #status experimental  
F:27-83,140-200,246-304/Disulfide bonds: #status experimental  
F:102,103,106,109/Disulfide bonds: interchain (to heavy chain) #status experimental  
F:176/Binding site: carbohydrate (Asn) (covalent) #status predicted  
Query Match 80.3%; Score 1148; DB 1; Length 326;  
Best Local Similarity 72.68; Pred. No. 1.4e-30;  
Matches 217; Conservative 13; Mismatches 20; Indels 49; Gaps 4;

QY 14 LKEAF-----DW-----LKAIFYDKVAEK 31  
Db 29 VKDYEPPEVTVSMNSGALTSGVHTFPAVLQSSGLYSLSSVTVPPSSNFQTQTYTCNVDRH 88  
QY 32 LKEAFMDKT-----HTCPPCPAPELLGSPSVFLFPKPDKDTLMISRPEVTCVVVDYSH 85  
Db 89 PSNTKVDKTVKRCCKVECPAPP-VAGPSVFLFPKPDKDTLMISRPEVTCVVVDYSH 147  
QY 86 EDPEVKFNWYVDGVVHNNAKTPREQVNSTYRVVSVLTVLHODWLNKKEYCKKYSNKAL 145  
Db 148 EDPEVQFNWYVDGVVHNNAKTPREQVNSTYRVVSVLTVLHODWLNKKEYCKKYSNKGL 207  
QY 146 PAPIEKTISKAGQPREQVYVTLPPSRDELTKNQVSLTCLVKGFYPSDIAVWESNGQPE 205  
Db 208 PAPIEKTISKTKGQPREQVYVTLPPSRDELTKNQVSLTCLVKGFYPSDIAVWESNGQPE 267  
QY 206 NNYKTPPVLDSDGSFFLYSKLTVDKSRWQQGNVFSCSVMHEALHNHYTQKSLSLSPGK 264  
Db 268 NNYKTPPMLDSDGSFFLYSKLTVDKSRWQQGNVFSCSVMHEALHNHYTQKSLSLSPGK 326  
RESULT 8  
GAHU  
Ig gamma-4 chain C region - human  
C:Species: Homo sapiens (man)  
C>Date: 02-Apr-1982 #sequence\_revision 02-Apr-1982 #text\_change 16-Jul-1999  
C:Accession: A90933; A90249; A02150  
R:Ellison, J.; Buxbaum, J.; Hood, L.  
DNA 1, 11-18, 1981  
A:Title: Nucleotide sequence of a human immunoglobulin C-gamma4 gene.  
A:Reference number: A90933; MUID:83157104; PMID:6299662  
A:Accession: A90933  
A:Molecule type: DNA  
A:Residues: 1-327 <ELL>  
A:Note: the sequence was determined from the germline gene  
R:Pink, J.R.L.; Buttery, S.H.; De Vries, G.M.; Milstein, C.  
Biochem. J. 117, 33-47, 1970  
A:Title: Human immunoglobulin subclases. Partial amino acid sequence of the constan  
A:Reference number: A90249; MUID:70207560; PMID:4192699  
A:Accession: A90249  
A:Molecule type: protein  
A:Residues: 1-30; 81-326 <PIN>  
C:Genetics:  
A:Gene: GDB:IGHG4  
A:CROSS-references: GDB:119340; OMIM:147130  
A:Map position: 14q32.33-14q32.33  
A:Introns: 99/1; 111/1; 221/1  
C:Complex: An immunoglobulin heterotetramer subunit consists of two identical light (C  
hain disulfide bonds. In some cases, such as IgA and IgM, the subunits associate into  
C:Superfamily: immunoglobulin C region; immunoglobulin homology  
C:Keywords: duplication; glycoprotein; heterotetramer; immunoglobulin  
F:20-85/Domain: immunoglobulin homology <IM1>  
F:99-110/Region: hinge  
F:134-203/Domain: immunoglobulin homology <IM2>  
F:240-307/Domain: immunoglobulin homology <IM3>  
F:14/Disulfide bonds: interchain (to light chain) #status experimental  
F:27-83,141-201,247-305/Disulfide bonds: #status predicted  
F:106,109/Disulfide bonds: interchain (to heavy chain) #status experimental  
F:177/Binding site: carbohydrate (Asn) (covalent) #status predicted  
Query Match 80.0%; Score 1144; DB 1; Length 327;  
Best Local Similarity 84.7%; Pred. No. 1.9e-30;  
Matches 211; Conservative 9; Mismatches 6; Indels 23; Gaps 1;  
QY 39 KTHT-----CPCPAPDELGGPSVFLFPKPDKDTLMISRTP 75  
Db 79 KTYTCNVDRHDKPSNTKVDKRVESKYGPCPCPAPEFLGGPSVFLFPKPDKDTLMISRTP 138  
QY 76 VTCVVVDVSHEDPEVKFNWYVDGVVHNNAKTPREQVNSTYRVVSVLTVLHODWLNK 135  
Db 139 VTCVVVDVSHEDPEVKFNWYVDGVVHNNAKTPREQVNSTYRVVSVLTVLHODWLNK 198  
QY 136 YKCKVSNKALPAPIEKTISKAKGQPREPVYVTLPPSRDELTKNQVSLTCLVKGFYPSDIA 195



F:173/Binding site: carbohydrate (Asn) (covalent) #status predicted

```
Query Match          64.5%; Score 923; DB 1; Length 323;
Best Local Similarity 69.8%; Pred. No. 4.5e-23;
Matches 169; Conservative 29; Mismatches 39; Indels 5; Gaps 2;

QY 28 VAEKLEAFMDKT---HTC--PPCPAPPELLGGPSVFLFPPKPKDTLMISRTPEVTCVVD 82
   |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
Db 82 VAHPATNTKVDKTVAPSTCKSPCPPPELLGGPSVFLFPPKPKDTLMISRTPEVTCVVD 141

QY 83 VSHEDPEVKFNWYDGVGVHNAKTKPREQYNSTYRVVSVLTVLHQDWLNGKEYKCKVSN 142
   |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
Db 142 VSDDPEVQFTTWINNEQVARTAPRPEEQFNSTIRWSTLPITHQDWLRGKFKCKVHN 201

QY 143 KALPAPIETKISKAKQPREPOVYTLPPSRDELTKNQVSLTCLVKGFYPSDIAVEWESNG 202
   |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
Db 202 KALPAPIETKISKAKQPREPKYVTMGPPREELSSRSVSLTCHINGFYPSDISVEWCKNG 261

QY 203 QPENNYKTTTPPVLDSDGSFLLSKLTVDKSRWQGNVFCSCVMHEALHNHYTKQSLSP 262
   |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
Db 262 KAEDNYKTTTPAVLDSDGSFLLSKLTVDKSRWQGNVFCSCVMHEALHNHYTKQSLSP 321

QY 263 GK 264
   ||
Db 322 GK 323
```

## RESULT 11

```
I47160
Ig gamma 2b chain constant region - pig (fragment)
C:Species: Sus scrofa domestica (domestic pig)
C>Date: 21-Feb-1997 #sequence_revision 21-Feb-1997 #text_change 21-Jan-2000
C:Accession: I47160
R:Kacskovics, I.; Sun, J.; Butler, J.E.
J. Immunol. 153, 3565-3573, 1994
A:Title: Five putative subclasses of swine IgG identified from the cDNA sequences of a
A:Reference number: I47158; MUID:95015845; PMID:7930579
A:Accession: I47160
A>Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-328 <KAC>
A:Cross-references: EMBL:U03780; NID:9433125; PIDN:AAA52218.1; PID:9433126
C:Genetics:
A:Gene: IgG2b
C:Superfamily: immunoglobulin C region; immunoglobulin homology
F:133-202/Domain: immunoglobulin homology <IMM>
```

```
Query Match          64.1%; Score 917; DB 2; Length 328;
Best Local Similarity 65.4%; Pred. No. 7.5e-23;
Matches 166; Conservative 32; Mismatches 24; Indels 32; Gaps 4;

QY 39 KTHTC-----PPCP-----APELLGGPSVFLFPPKPKDTLMISR 72
   |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
Db 79 KSYTCNVNHPATTTKVKRVGKTKPPCPICPACESP-----GPSVFIFPPKPKDTLMISR 134

QY 73 TPEVTCVVVDVSHEDPEVKFNWYDGVGVHNAKTKPREQYNSTYRVVSVLTVLHQDWLN 132
   |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
Db 135 TPQVTCVVVDVSOENPEVQFSWYDGVGVHNAKTKPREQYNSTYRVVSVLTVLHQDWLN 194

QY 133 GREYCKVKSNKALPAPIETKISKAKQPREPOVYTLPPSRDELTKNQVSLTCLVKGFYPS 192
   |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
Db 195 GREYCKVKSNKALPAPIETKISKAKQPREPOVYTLPPSRDELTKNQVSLTCLVKGFYPS 254

QY 193 DIAVWESNGQ--PENNYKTTTPVLDSDGSFLLSKLTVDKSRWQGNVFCSCVMHEALH 250
   |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
Db 255 DIDVWQRNGQPEPEGNYYTTPQDDVDGYLYSKFSVDKASWGGGIFQCAVMHEALH 314

QY 251 NHYTKQSLSPGK 264
   |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
Db 315 NHYTKQSLSPGK 328
```

## RESULT 12

```
I47159
Ig gamma 2a chain constant region - pig (fragment)
C:Species: Sus scrofa domestica (domestic pig)
C>Date: 21-Feb-1997 #sequence_revision 21-Feb-1997 #text_change 21-Jan-2000
C:Accession: I47159
R:Kacskovics, I.; Sun, J.; Butler, J.E.
J. Immunol. 153, 3565-3573, 1994
A:Title: Five putative subclasses of swine IgG identified from the cDNA sequences of
A:Reference number: I47158; MUID:95015845; PMID:7930579
A:Accession: I47159
A>Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-277 <KAC>
A:Cross-references: EMBL:U03782; NID:9433129; PIDN:AAA52220.1; PID:9433130
C:Genetics:
A:Gene: IgG4
C:Superfamily: immunoglobulin C region; immunoglobulin homology
F:82-151/Domain: immunoglobulin homology <IMM>
```

```
I47159
Ig gamma 2a chain constant region - pig (fragment)
C:Species: Sus scrofa domestica (domestic pig)
C>Date: 21-Feb-1997 #sequence_revision 21-Feb-1997 #text_change 21-Jan-2000
C:Accession: I47159
R:Kacskovics, I.; Sun, J.; Butler, J.E.
J. Immunol. 153, 3565-3573, 1994
A:Title: Five putative subclasses of swine IgG identified from the cDNA sequences of
A:Reference number: I47158; MUID:95015845; PMID:7930579
A:Accession: I47159
A>Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-328 <KAC>
A:Cross-references: EMBL:U03779; NID:9433123; PIDN:AAA52217.1; PID:9433124
C:Genetics:
A:Gene: IgG2a
C:Superfamily: immunoglobulin C region; immunoglobulin homology
F:133-202/Domain: immunoglobulin homology <IMM>
```

```
Query Match          64.1%; Score 917; DB 2; Length 328;
Best Local Similarity 65.4%; Pred. No. 7.5e-23;
Matches 166; Conservative 32; Mismatches 24; Indels 32; Gaps 4;

QY 39 KTHTC-----PPCP-----APELLGGPSVFLFPPKPKDTLMISR 72
   |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
Db 79 KSYTCNVNHPATTTKVKRVGKTKPPCPICPACESP-----GPSVFIFPPKPKDTLMISR 134

QY 73 TPEVTCVVVDVSHEDPEVKFNWYDGVGVHNAKTKPREQYNSTYRVVSVLTVLHQDWLN 132
   |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
Db 135 TPQVTCVVVDVSOENPEVQFSWYDGVGVHNAKTKPREQYNSTYRVVSVLTVLHQDWLN 194

QY 133 GREYCKVKSNKALPAPIETKISKAKQPREPOVYTLPPSRDELTKNQVSLTCLVKGFYPS 192
   |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
Db 195 GREYCKVKSNKALPAPIETKISKAKQPREPOVYTLPPSRDELTKNQVSLTCLVKGFYPS 254

QY 193 DIAVWESNGQ--PENNYKTTTPVLDSDGSFLLSKLTVDKSRWQGNVFCSCVMHEALH 250
   |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
Db 255 DIDVWQRNGQPEPEGNYYTTPQDDVDGYLYSKFSVDKASWGGGIFQCAVMHEALH 314

QY 251 NHYTKQSLSPGK 264
   |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
Db 315 NHYTKQSLSPGK 328
```

## RESULT 13

```
I47162
Ig gamma 4 chain constant region - pig (fragment)
C:Species: Sus scrofa domestica (domestic pig)
C>Date: 21-Feb-1997 #sequence_revision 21-Feb-1997 #text_change 21-Jan-2000
C:Accession: I47162
R:Kacskovics, I.; Sun, J.; Butler, J.E.
J. Immunol. 153, 3565-3573, 1994
A:Title: Five putative subclasses of swine IgG identified from the cDNA sequences of
A:Reference number: I47158; MUID:95015845; PMID:7930579
A:Accession: I47162
A>Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-277 <KAC>
A:Cross-references: EMBL:U03782; NID:9433129; PIDN:AAA52220.1; PID:9433130
C:Genetics:
A:Gene: IgG4
C:Superfamily: immunoglobulin C region; immunoglobulin homology
F:82-151/Domain: immunoglobulin homology <IMM>
```

```
Query Match          63.7%; Score 911; DB 2; Length 277;
Best Local Similarity 66.4%; Pred. No. 6.8e-23;
Matches 166; Conservative 31; Mismatches 29; Indels 24; Gaps 4;

QY 39 KTHTC-----PPCP-----APELLG-GPSVFLFPPKPKDTLMISRTPEV 76
   |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
Db 28 KSYTCNVNHPATTTKVKRVGKTKPPCPICPACESPGPSAFIFPPKPKDTLMISRTPKV 87

QY 77 TCVVVDVSHEDPEVKFNWYDGVGVHNAKTKPREQYNSTYRVVSVLTVLHQDWLNGKEY 136
```

Search completed: April 21, 2003, 10:49:19  
Job time : 20.4118 secs





GenCore version 5.1.4\_p5\_4578  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: April 21, 2003, 10:42:21 ; Search time 10.0941 Seconds  
(without alignments)  
1084.766 Million cell updates/sec

Title: 7LINK7LINK2

Perfect score: 1430

Sequence: 1 DMLKAFYDKVAEKLKAPFW.....MHEALHNHYTKLSLSPGK 264

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.0

Searched: 112892 seqs, 41476328 residues

Total number of hits satisfying chosen parameters: 112892

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt\_40:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match %	Length	ID	Description
1	1243	86.9	330	1 GC1_HUMAN	P01857 homo sapien
2	1148	80.3	290	1 GC3_HUMAN	P01860 homo sapien
3	1148	80.3	326	1 GC2_HUMAN	P01859 homo sapien
4	1144	80.0	327	1 GC4_HUMAN	P01861 homo sapien
5	923	64.5	323	1 GC_RABIT	P01870 oryctolagus
6	899	62.9	329	1 GC2_CAVPO	P01862 cavia porce
7	853	59.7	329	1 GC3_MOUSE	P22436 mus musculu
8	849	59.4	333	1 GCB_RAT	P20761 rattus norv
9	842	58.9	398	1 GC3M_MOUSE	P03987 mus musculu
10	826	57.8	326	1 GCL_RAT	P20759 rattus norv
11	823	57.6	324	1 GCL_MOUSE	P01868 mus musculu
12	818	57.2	329	1 GCC_RAT	P20762 rattus norv
13	818	57.2	393	1 GC1M_MOUSE	P01869 mus musculu
14	815	57.0	330	1 GCAA_MOUSE	P01863 mus musculu
15	810	56.6	399	1 GCAM_MOUSE	P01865 mus musculu
16	809	56.6	335	1 GCAB_MOUSE	P01864 mus musculu
17	795	55.6	322	1 GCA_RAT	P20760 rattus norv
18	789	55.2	336	1 GCB_MOUSE	P01866 mus musculu
19	784	54.8	405	1 GCBM_MOUSE	P01867 mus musculu
20	389	27.2	421	1 EPC_MOUSE	P06336 mus musculu
21	386	27.0	455	1 MUC_MOUSE	P01872 mus musculu
22	383	26.8	454	1 MUC_HUMAN	P01871 homo sapien
23	382	26.7	428	1 EPC_HUMAN	P01854 homo sapien
24	381	26.6	391	1 MUCB_HUMAN	P04220 homo sapien
25	378	26.4	429	1 EPC_RAT	P01855 rattus norv
26	376	26.3	458	1 MUC_RABIT	P03988 oryctolagus
27	376	26.3	476	1 MUCM_MOUSE	P01873 mus musculu
28	366	25.6	479	1 MUCM_RABIT	P04221 oryctolagus
29	363	25.4	454	1 MUC_MESAU	P06337 mesocricetu
30	356	24.9	457	1 MUC_SUNMU	P20768 suncus muri
31	354	24.8	450	1 MUC_CANFA	P01874 canis fami
32	343	24.0	438	1 HVC2_HETER	P23085 heterodontu
33	335	23.4	438	1 HVC3_HETER	P23087 heterodontu

RESULT 1					ALIGNMENTS				
ID	GC1_HUMAN	STANDARD;	PRT;	330 AA.					
AC	P01857;								
DT	21-JUL-1986 (Rel. 01, Created)								P20929 homo sapien
DT	21-JUL-1986 (Rel. 01, Last sequence update)								P23088 heterodontu
DT	16-OCT-2001 (Rel. 40, Last annotation update)								P01875 gallus gall
DE	Ig gamma-1 chain C region.								O68008 b bacitraci
GN	IGHG1.								O30409 b tyrocidin
OS	Homo sapiens (Human).								P23086 heterodontu
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;								O92736 homo sapien
OC	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.								P30957 oryctolagus
OX	NCBI_TaxID=9606;								Q39565 chlamydomon
RN	[1]								P23084 heterodontu
RP	SEQUENCE FROM N.A.								P23098 tripneustes
RX	MEDLINE=82274238; PubMed=6287432;								P98160 homo sapien
RA	Ellison J.W., Berson B.J., Hood L.E.;								
RT	"The nucleotide sequence of a human immunoglobulin C gamma1 gene.";								
RL	Nucleic Acids Res. 10:4071-4079(1982).								
RN	[2]								
RP	SEQUENCE OF 1-135 (MYELOMA PROTEIN EU).								
RX	MEDLINE=71064024; PubMed=5489771;								
RA	Cunningham B.A., Rutishauser U., Gall W.E., Gottlieb P.D.,								
RT	Waxdal M.J., Edelman G.M.;								
RL	"The covalent structure of a human gamma G-immunoglobulin. VII. Amino acid sequence of heavy-chain cyanogen bromide fragments H1-H4.";								
RN	[3]								
RP	Biochemistry 9:3161-3170(1970).								
RX	SEQUENCE OF 136-329 (EU).								
RA	MEDLINE=71064025; PubMed=5530842;								
RT	Rutishauser U., Cunningham B.A., Bennett C., Konigsberg W.H.,								
RL	Edelman G.M.;								
RT	"The covalent structure of a human gamma G-immunoglobulin. 8. Amino acid sequence of heavy-chain cyanogen bromide fragments H5-H7.";								
RN	[4]								
RP	Biochemistry 9:3171-3181(1970).								
RX	SEQUENCE (MYELOMA PROTEIN NIE).								
RA	MEDLINE=77070269; PubMed=826475;								
RT	Ponstingl H., Hilschmann N.;								
RL	"The rule of antibody structure. The primary structure of a monoclonal IgG1 immunoglobulin (myeloma protein Nie). III. The chymotryptic peptides of the H-chain, alignment of the tryptic peptides and discussion of the complete structure.";								
RN	[5]								
RP	Hoppe-Seyler's Z. Physiol. Chem. 357:1571-1604(1976).								
RX	SEQUENCE (MYELOMA PROTEIN KOL), AND DISULFIDE BONDS.								
RA	MEDLINE=83289131; PubMed=6884994;								
RT	Schmidt W.E., Jung H.-D., Palm W., Hilschmann N.;								
RL	"Three-dimensional structure determination of antibodies. Primary structure of crystallized monoclonal immunoglobulin IgG1 KOL, I.";								
RN	[6]								
RP	Hoppe-Seyler's Z. Physiol. Chem. 364:713-747(1983).								
RX	DISULFIDE BONDS.								
RA	MEDLINE=71064027; PubMed=4923144;								
RT	Gall W.E., Edelman G.M.;								
RL	"The covalent structure of a human gamma G-immunoglobulin. X.								

RT Intrachain disulfide bonds.";  
 RL Biochemistry 9:3188-3196(1970).  
 RN [7]  
 RP DISULFIDE BONDS.  
 RA MEDLINE=77070267; PubMed=1002129;  
 RX Dreker L., Schwarz J., Reichel W., Hilschmann N.;  
 RT "Rule of antibody structure. The primary structure of a monoclonal  
 RT IgG1 immunoglobulin (myeloma protein Nie), I: Purification and  
 RT characterization of the protein, the L- and H-chains, the  
 RT cyanogen bromide cleavage products, and the disulfide bridges.";  
 RL Hoppe-Seyler's Z. Physiol. Chem. 357:1515-1540(1976).  
 RN [8]  
 RP X-RAY CRYSTALLOGRAPHY (2.9 ANGSTROMS).  
 RA MEDLINE=81208100; PubMed=7236608;  
 RX Deisenhofer J.;  
 RT "Crystallographic refinement and atomic models of a human Fc fragment  
 RT aureus at 2.9- and 2.8-A resolution.";  
 RL Biochemistry 20:2361-2370(1981).  
 CC -1- MISCELLANEOUS: NIE HAS THE GIM(17) ALLOTYPIC MARKER, 97-K, & THE  
 CC GIM(1) MARKERS, 239-D & 241-L. KOL & EU SEQUENCES HAVE THE GIM(3)  
 CC MARKER & THE GIM (NON-1) MARKERS.  
 CC -1- MISCELLANEOUS: NIE ALSO DIFFERS IN THE AMIDATION STATES OF  
 CC 35,116,198,269 & 272.  
 CC -1- MISCELLANEOUS: EU ALSO DIFFERS IN THE AMIDATION STATES OF RESIDUES  
 CC 155, 166, 177, 195, 198, 269, AND 272 AND IN THE ORDER OF RESIDUES  
 CC 268-272.  
 CC -1- MISCELLANEOUS: KOL ALSO DIFFERS IN THE AMIDATION STATES OF  
 CC RESIDUES 198,267&272.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (see <http://www.isb-sib.ch/Announce/>  
 CC or send an email to [license@sib-sib.ch](mailto:license@sib-sib.ch)).  
 CC -----  
 DR EMBL; J00228; AAC82527.1; ALT\_INIT.  
 DR PIR; A02146; GHU.  
 DR PDB; 1FC1; 15-JUL-92.  
 DR PDB; 1FC2; 15-JUL-92.  
 DR Genew; HGNC:5525; IGHG1.  
 DR MIM; 147100; -  
 DR InterPro; IPR003306; Ig\_MHC.  
 DR InterPro; IPR003597; Ig\_c1.  
 DR InterPro; IPR003600; Ig\_like.  
 DR Pfam; PF00047; Ig; 3.  
 DR SMART; SM00410; Ig\_like; 1.  
 DR SMART; SM00407; IGHG1; 2.  
 DR PROSITE; PS00290; Ig\_MHC; 2.  
 KW Immunoglobulin domain; Immunoglobulin C region; Glycoprotein;  
 KW 3D-structure.  
 FT NON\_TER 1 1  
 FT DOMAIN 1 98  
 FT CH1.  
 FT HINGE.  
 FT 99 110  
 FT CH2.  
 FT 111 223  
 FT CH3.  
 FT 224 330  
 FT DISULFID 27 83  
 FT DISULFID 103 103  
 FT DISULFID 109 109  
 FT DISULFID 112 112  
 FT DISULFID 144 204  
 FT DISULFID 250 308  
 FT CARBOHYD 180 180  
 FT MOD\_RES 330 330  
 FT VARIANT 97 97  
 FT 239 239  
 FT VARIANT 241 241  
 FT VARIANT 241 241  
 FT STRAND 123 126

FT HELIX 130 134  
 FT TURN 136 137  
 FT STRAND 141 148  
 FT STRAND 158 162  
 FT TURN 163 164  
 FT STRAND 165 166  
 FT STRAND 175 178  
 FT STRAND 183 190  
 FT STRAND 193 197  
 FT HELIX 198 199  
 FT TURN 202 206  
 FT STRAND 215 219  
 FT STRAND 227 227  
 FT STRAND 230 234  
 FT STRAND 238 240  
 FT TURN 241 242  
 FT STRAND 245 256  
 FT STRAND 260 266  
 FT TURN 267 268  
 FT STRAND 274 276  
 FT STRAND 280 281  
 FT TURN 283 284  
 FT STRAND 287 296  
 FT HELIX 297 301  
 FT TURN 302 303  
 FT STRAND 306 312  
 FT TURN 313 314  
 FT TURN 316 317  
 FT STRAND 320 324  
 SQ SEQUENCE 330 AA; 36106 MW; 3770EE106C2FA33D CRC64;  
 Query Match 86.98; Score 1243; DB 1; Length 330;  
 Best Local Similarity 96.68; Pred. No. 1.3e-34;  
 Matches 230; Conservative 3; Mismatches 5; Indels 0; Gaps 0;  
 QY 27 KVAEKLKAEAFMDKTHTCPPCPAPPELLGGPSVFLFPPPKPDTLMISRTPEVTCVVVDVSH 86  
 DB 93 KVDKKVEPKSCDKTHTCPPCPAPPELLGGPSVFLFPPPKPDTLMISRTPEVTCVVVDVSH 152  
 QY 87 DPEVKFNWYDGVVHNAKTPREEQYNSTYRVVSVLTVLHQDWLNGKEYKCKVSNKALP 146  
 DB 153 DPEVKFNWYDGVVHNAKTPREEQYNSTYRVVSVLTVLHQDWLNGKEYKCKVSNKALP 212  
 QY 147 APIEKTISKAKGQPREPOVYTLPPSRDELTKNQVSLTCLVKGFYPSDIAVWESNGQPEN 206  
 DB 213 APIEKTISKAKGQPREPOVYTLPPSRDELTKNQVSLTCLVKGFYPSDIAVWESNGQPEN 272  
 QY 207 NYKTTTPVLDSDGSFFLYSKLTVDKSRWQQGNVSCFVSVMHEALHNHYTQKSLSLSPGK 264  
 DB 273 NYKTTTPVLDSDGSFFLYSKLTVDKSRWQQGNVSCFVSVMHEALHNHYTQKSLSLSPGK 330  
 RESULT 2  
 GC3\_HUMAN  
 ID GC3\_HUMAN STANDARD; PRT; 230 AA.  
 AC P01860;  
 DT 21-JUL-1986 (Rel. 01, Created)  
 DT 21-JUL-1986 (Rel. 01, Last sequence update)  
 DE 15-JUN-2002 (Rel. 41, Last annotation update)  
 GN IGHG3.  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
 OX NCBI\_TaxID=9606;  
 RN [1]  
 RP SEQUENCE (DISEASE PROTEIN WIS).  
 RX MEDLINE=81021548; PubMed=6774747;  
 RA Frangione B., Rosenwasser E., Prelli F., Franklin E.C.;  
 RT "Primary structure of human gamma 3 immunoglobulin deletion mutant:  
 RT gamma 3 heavy-chain disease protein wis.";  
 RL Biochemistry 19:4304-4308(1980).

FT	DISULFID	39	INTERCHAIN (WITH HEAVY CHAIN DIMER).
FT	DISULFID	42	INTERCHAIN (WITH HEAVY CHAIN DIMER).
FT	DISULFID	48	INTERCHAIN (WITH HEAVY CHAIN DIMER).
FT	DISULFID	54	INTERCHAIN (WITH HEAVY CHAIN DIMER).
FT	DISULFID	57	INTERCHAIN (WITH HEAVY CHAIN DIMER).
FT	DISULFID	63	INTERCHAIN (WITH HEAVY CHAIN DIMER).
FT	DISULFID	69	INTERCHAIN (WITH HEAVY CHAIN DIMER).
FT	DISULFID	72	INTERCHAIN (WITH HEAVY CHAIN DIMER).
FT	CARBOHYD	140	INTERCHAIN (WITH HEAVY CHAIN DIMER).
FT	MOD_RES	290	N-LINKED (GLCNAC..).
FT	VARIANT	126	REMOVED POST-TRANSLATIONALLY.
FT	VARIANT	134	QV -> EB (IN ZUC).
FT	VARIANT	139	/FTID=VAR_003890.
FT	VARIANT	182	P -> L (IN OHM).
FT	VARIANT	227	/FTID=VAR_003891.
FT	VARIANT	227	F -> Y (IN OHM).
FT	VARIANT	227	T -> A (IN OHM).
FT	VARIANT	227	S -> N (IN OHM).
FT	VARIANT	227	/FTID=VAR_003894.
FT	VARIANT	279	MISSING (IN ZUC).
FT	VARIANT	279	/FTID=VAR_003895.
FT	VARIANT	279	F -> Y (IN OHM).
FT	VARIANT	279	/FTID=VAR_003896.
FT	SEQUENCE	290 AA;	E69CBC95705B2F46 CRC64;
FT	SEQUENCE	290 AA;	32331 MW; E69CBC95705B2F46 CRC64;
QY	Query Match	80.3%;	Score 1148; DB 1; Length 290;
QY	Best Local Similarity	74.8%;	Pred. No. 1.5e-31;
DB	Matches 211;	Conservative 15;	Mismatches 11; Indels 45; Gaps
QY	28 VAEKLKEAFMDKTHVT-----	42	
DB	: : : :		
QY	9 VSELTKTPLGDTHTTCPCPEPKSCDTPPPCPCPEPKSCDTPPPCPCPEPKSCDTPPP	68	
QY	43 CPCCPAPELLGGPSVFLFPKPKPDKTLMISRTPEVTCVVVDVSHEDPEVKFNWYVDGV	102	
DB	: : : :		
QY	69 CPCRPAPELLGGPSVFLFPKPKPDKTLMISRTPEVTCVVVDVSHEDPEVKFNWYVDGV	128	
QY	103 NAKTKPREEQNYSTYRVSVLVTLVHQLDWLNGKEYCKVSNKALPAPIEKTISKAKGQPR	162	
DB	: : : :		
QY	129 NAKTKPREQQNFSTFRVSVLTVLHQMLDGKEYCKVSNKALPAPIEKTISKAKGQPR	188	
QY	163 PQVYTLPSPRDELTKNOVSLTCLVKGFPSPDIADVESNGQPENNYKTTPPVLDSDGSFF	222	
DB	: : : :		
QY	189 PQVYTLPSPRDEMTKNQVSLTCLVKGFPSPDIADVESHSSGPENNTTPPMLDSGDSFF	248	
QY	223 LYSKLTVDKSRWQGNVFSCVMHEALHNHYTKLSLSPGK	264	
DB	: : : :		
QY	249 LYSKLTVDKSRWQGNIFSCVMHEALHNRFQKLSLSPGK	290	
RESULT 3			
ID	GC2_HUMAN	STANDARD;	PRT; 326 AA.
AC	P01859;		
DC	21-JUL-1986 (Rel. 01, Created)		
DT	21-JUL-1986 (Rel. 01, Last sequence update)		
DE	16-OCT-2001 (Rel. 40, Last annotation update)		
OS	Ig gamma-2 chain c region.		
GN	IGHG2.		
OC	Homo sapiens (Human).		
OC	Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;		
OC	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.		
OX	NCBI_TaxID=9606;		
RN	[1]		
RP	SEQUENCE OF 2-326 FROM N.A.		
RA	MEDLINE=82197621; PubMed=6804948;		
RX	Ellison J W., Hood L.E.;		
RT	"Linkage and sequence homology of two human immunoglobulin gamma		
RL	heavy chain constant region genes.";		
RL	Proc. Natl. Acad. Sci. U.S.A. 79:1984-1988(1982).		
RP	[2]		
RP	SEQUENCE OF 88-115 FROM N.A.		

RC TISSUE-Fetal liver;  
RX MEDLINE=83001943; PubMed=6811139;  
RA Takahashi N., Ueda S., Obata M., Nikaigo T., Nakai S., Honjo, T.;  
RT "Structure of human immunoglobulin gamma genes: implications for  
evolution of a gene family.";  
RL Cell 29:671-679(1982).  
RN [13]  
RP SEQUENCE OF 99-177 AND 310-326 FROM N.A.  
RC TISSUE-Fetal liver;  
RX MEDLINE=84235992; PubMed=6329676;  
RA Krawinkel U., Rabbits T.H.;  
RT "Comparison of the hinge-coding segments in human immunoglobulin gamma  
heavy chain genes and the linkage of the gamma 2 and gamma 4 subclass  
genes.";  
RL EMBO J. 1:403-407(1982).  
RN [4]  
RP SEQUENCE OF 1-325 (MYELOMA PROTEIN TIL).  
RX MEDLINE=81007873; PubMed=6774012;  
RA Wang A.-C., Tung E., Fudenberg H.H.;  
RT "The primary structure of a human IgG2 heavy chain: genetic,  
evolutionary, and functional implications.";  
RL J. Immunol. 125:1048-1054(1980).  
RN [5]  
RP SEQUENCE OF 1-85 AND 132-325 (MYELOMA PROTEIN ZIE).  
RX MEDLINE=80001357; PubMed=113060;  
RA Connell G.E., Parr D.M., Hofmann T.;  
RT "The amino acid sequences of the three heavy chain constant region  
domains of a human IgG2 myeloma protein.";  
RL Can. J. Biochem. 57:758-767(1979).  
RN [6]  
RP SEQUENCE OF 238-275 (ZIE).  
RX MEDLINE=80114419; PubMed=118920;  
RA Hofmann T., Parr D.M.;  
RT "A note of the amino acid sequence of residues 381-391 of human  
immunoglobulin gamma chains.";  
RL Mol. Immunol. 16:923-925(1979).  
RN [7]  
RP REVISIONS TO 25; 59; 60 AND 264-268 (ZIE).  
RA Hofmann T., Parr D.M.;  
RT Submitted (MAR-1980) to the PIR data bank.  
RN [8]  
RP SEQUENCE OF 1-121 (DOT).  
RX MEDLINE=93255298; PubMed=7737190;  
RA Stoppini M., Bellotti V., Negri A., Merlini G., Garver F., Ferri G.;  
RT "Characterization of the two unique human anti-flavon monoclinal  
immunoglobulins.";  
RL Eur. J. Biochem. 228:886-893(1995).  
RN [9]  
RP DISULFIDE BONDS.  
RX MEDLINE=72033500; PubMed=4940472;  
RA Milstein C., Frangione B.;  
RT "Disulphide bridges of the heavy chain of human immunoglobulin G2.";  
RL Biochem. J. 121:217-225(1971).  
RN [10]  
RP DISULFIDE BONDS.  
RX MEDLINE=69064124; PubMed=5782707;  
RA Frangione B., Milstein C., Pink J.R.L.;  
RT "Structural studies of immunoglobulin G.";  
RL Nature 221:145-148(1969).  
RN [11]  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
between the Swiss Institute of Bioinformatics and the EMBL Outstation -  
the European Bioinformatics Institute. There are no restrictions on its  
use by non-profit institutions as long as its content is in no way  
modified and this statement is not removed. Usage by and for commercial  
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
RN [12]  
DR EMBL: J00230; AAB59393.1; -  
RX A02148; G2HU.  
DR HSSP: P01857; 1FC1.  
DR Genew; HGNC:5526; IGHG2.  
RN 147110; -

DR InterPro; IPR003006; Ig\_MHC.  
DR InterPro; IPR003597; Ig\_c1.  
DR InterPro; IPR003600; Ig\_like.  
DR Pfam; PF00047; ig; 3.  
DR SMART; SM00410; Ig\_like; 1.  
DR SMART; SM00407; IGC1; 2.  
DR PROSITE; PS00290; IG\_MHC; 2.  
KW Immunoglobulin domain; Immunoglobulin C region.  
FT NON\_TER 1 1  
FT DOMAIN 1 98  
FT HINGE. 110  
FT DOMAIN 111 219  
FT DOMAIN 220 326  
FT DISULFID 14 14  
FT DISULFID 27 83  
FT DISULFID 102 102  
FT DISULFID 103 103  
FT DISULFID 106 106  
FT DISULFID 109 109  
FT DISULFID 140 200  
FT DISULFID 246 304  
FT SITE 156 156  
FT MOD\_RES 326 326  
FT VARIANT 60 60  
FT CONFLICT 109 109  
FT SEQUENCE 326 AA; 35884 MW; 8310878C6878CF9C CRC64;  
SQ  
Query Match 80.3%; Score 1148; DB 1; Length 326;  
Best Local Similarity 72.6%; Pred. No. 2.le-31;  
Matches 217; Conservative 13; Mismatches 20; Indels 49; Gaps 4;  
QY 14 LKEAF-----DW-----LKAFLDKVAKK 31  
DB 29 VKDYPEPVTVSWNSGALTSGVHTFPVAVLQSSGLYSLSSVTVVPSNFTQYTCNVDHK 88  
QY 32 LKEAFMDKT-----HTCPPCPAPPELLGGPSVFLFPPPKPKDTLMISRTPEVTCVVVDVSH 85  
DB 89 PSNTRVDTKVERKCCVCEPCPAPP-VAGPSVFLFPPPKPKDTLMISRTPEVTCVVVDVSH 147  
QY 86 EDPEVKFNWYDGVGVHNAKTPREEQYNSTYRVVSVLTVLHQDWLNGKEYCKCKVSNKAL 145  
DB 148 EDPEVQFNWYDGVGVHNAKTPREEQYNSTYRVVSVLTVLHQDWLNGKEYCKCKVSNKGL 207  
QY 146 PAPIEKTISKAKGQPREPOVYTLPPSRDELTKNQVSLTCLVKGFYPSDIAVWESNGQPE 205  
DB 208 PAPIEKTISKAKGQPREPOVYTLPPSRDEMTKNQVSLTCLVKGFYPSDIAVWESNGQPE 267  
QY 206 NNYKTPPVLDSDGSFFLYSKLTVDKSRWQQGNVSCSVMHRLNNHHTQKLSLSLSPGK 264  
DB 268 NNYKTPPVLDSDGSFFLYSKLTVDKSRWQQGNVSCSVMHRLNNHHTQKLSLSLSPGK 326  
RESULT 4  
GC4\_HUMAN STANDARD; PRT; 327 AA.  
ID GC4\_HUMAN  
AC P01861.  
DT 21-JUL-1986 (Rel. 01, Created)  
DT 21-JUL-1986 (Rel. 01, Last sequence update)  
DT 16-OCT-2001 (Rel. 40, Last annotation update)  
DE Ig gamma-4 chain C region.  
GN IGHG4.  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
OX NCBI\_TaxID=9606;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=83157104; PubMed=6299662;  
RA Ellison J.W., Buxbaum J.N., Hood L.E.;  
RT "Nucleotide sequence of a human Immunoglobulin C gamma 4 gene.";  
RL DNA 1:11-18(1981).  
RN [2]

```
RP SEQUENCE OF 1-30 AND 81-326.
RX MEDLINE=70207560; PubMed=4192699;
RA Pink J.R.L., Buttery S.H., de Vries G.M., Milstein C.;
RT "Human immunoglobulin subclasses. Partial amino acid sequence of the
RL constant region of a gamma 4 chain.";
RL Biochem. J. 117:33-47(1970).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: K01316; AAB59394.1; ALT_INIT.
DR PIR: A02150; G4HU.
DR HSP: P01842; 7FAB.
DR Genew; HGNC:5528; IGHG4.
DR MIM; 147130;
DR InterPro: IPR003006; Ig_MHC.
DR InterPro: IPR003597; Ig_cl.
DR InterPro: IPR003600; Ig_like.
DR Pfam: PF00047; Ig; 3.
DR SMART; SM00410; Ig_like; 1.
DR SMART; SM00407; Igc1; 2.
DR PROSITE; PS00290; IG_MHC; 2.
KW Immunoglobulin domain; Immunoglobulin C region.
FT NON_TER 1
FT DOMAIN 1 98 CH1.
FT DOMAIN 99 110 HINGE.
FT DOMAIN 111 220 CH2.
FT DOMAIN 221 327 CH3.
FT DISULFID 14 14 INTERCHAIN (WITH A LIGHT CHAIN).
FT DISULFID 27 83
FT DISULFID 106 106 INTERCHAIN (WITH A HEAVY CHAIN).
FT DISULFID 109 109 INTERCHAIN (WITH A HEAVY CHAIN).
FT DISULFID 141 201
FT DISULFID 247 305
SQ SEQUENCE 327 AA; 35940 MW; 3EDBD811EF208E7A CRC64;

Query Match 80.0%; Score 1144; DB 1; Length 327;
Best Local Similarity 84.7%; Pred. No. 2.9e-31;
Matches 211; Conservative 9; Mismatches 6; Indels 23; Gaps 1;

QY 39 KTHF-----CPPCAPELLGSPVFLFPKPKDTLMISRTPE 75
Db 79 KTYCNVDHPSNTKVDKVESKYGPPCPAPFLGGPSVFLFPKPKDTLMISRTPE 138
QY 76 VTCVVDVSHDEPVEKFNWYVDGVEVHNATKPREQYNSTYRVSVTLVHODWLNKRE 135
Db 139 VTCVVDVDSQEDPEVQFNWYVDGVEVHNATKPREQFNSTYRVSVTLVHODWLNKRE 198
QY 136 YKCVSNKALPAPLEKTIKSAKGOPREPOVYTLPPSRDELTKNOVSLTCLVKGYPSDIA 195
Db 199 YKCVSNKGLPSLEKTIKSAKGOPREPOVYTLPPSQEEMTKNOVSLTCLVKGYPSDIA 258
QY 196 VEWESNGOPENNYKTTTPVLDSGFFLYSKLTVDKSRWQGNVFSQSVMEALHNHYTQ 255
Db 259 VEWESNGOPENNYKTTTPVLDSGFFLYSKLTVDKSRWQGNVFSQSVMEALHNHYTQ 318
QY 256 KSLSLSPCK 264
Db 319 KSLSLSLGK 327

RESULT 5
GC_RABIT
ID GC_RABIT STANDARD; PRT; 323 AA.
AC P01870;
DT 21-JUL-1986 (Rel. 01, Created)
DT 21-JUL-1986 (Rel. 01, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
```

```
DE Ig gamma chain C region.
OC Oryctolagus cuniculus (Rabbit).
OS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Lagomorpha; Leporidae; Oryctolagus.
OX NCBI_TaxID=9986;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=84030930; PubMed=6313520;
RA Bernstein K.E., Alexander C.B., Mage R.G.;
RT "Nucleotide sequence of a rabbit IgG heavy chain from the recombinant
RL F-1 haplotype.";
RL Immunogenetics 18:387-397(1983).
RN [2]
RP SEQUENCE OF 1-128.
RX MEDLINE=76135469; PubMed=1243651;
RA Pratt D.M., Moie L.E.;
RT "Sequence studies on the constant region of the Fd sections of rabbit
RL immunoglobulin G of different allotype.";
RL Biochem. J. 151:337-349(1975).
RN [3]
RP SEQUENCE OF 88-266 FROM N.A.
RX MEDLINE=83299917; PubMed=6193512;
RA Martens C.L., Moore K.W., Steinmetz M., Hood L., Knight K.L.;
RT "Heavy chain genes of rabbit IgG: isolation of a cDNA encoding gamma
RL heavy chain and identification of two genomic C gamma genes.";
RL Proc. Natl. Acad. Sci. U.S.A. 79:6018-6022(1982).
RN [4]
RP SEQUENCE OF 132-161.
RX MEDLINE=70110015; PubMed=5461106;
RA Fruchter R.G., Jackson S.A., Moie L.E., Porter R.R.;
RT "Sequence studies of the Fd section of the heavy chain of rabbit
RL immunoglobulin G.";
RL Biochem. J. 116:249-259(1970).
RN [5]
RP SEQUENCE OF 129-131 AND 155-322.
RA Hill R.L., Lebovitz H.E., Fellows R.E. Jr., Delaney R.;
RL (In) Killander J. (eds.);
RL Gamma globulins, Nobel symp. 3, pp.109-127, Almqvist and Wiksell,
RL Stockholm (1967).
CC -1- MISCELLANEOUS: REF.1 SEQUENCE HAS THE D12 ALLOTYPIC MARKER,
CC 104-THR, AND THE E14 MARKER, 185-THR. REF.3 HAS THE D11 AND E15
CC MARKERS AND REF.5 THE E15 MARKER.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: M16426; AAA31289.1; -.
DR PIR: A02161; GHRB.
DR HSP: P01857; IFC1.
DR InterPro: IPR003006; Ig_MHC.
DR InterPro: IPR003597; Ig_cl.
DR Pfam: PF00047; Ig; 2.
DR SMART; SM00407; Igc1; 2.
DR PROSITE; PS00290; IG_MHC; 1.
KW Immunoglobulin domain; Immunoglobulin C region.
FT NON_TER 1
FT VARIANT 104 104 T -> M (IN D11 MARKER).
FT VARIANT 185 185 T -> A (IN E15 MARKER).
FT CONFLICT 48 48 N -> E (IN REF. 2).
FT CONFLICT 71 71 V -> VPV (IN REF. 2).
FT CONFLICT 144 144 Q -> E (IN REF. 3 AND 4).
FT CONFLICT 173 173 N -> D (IN REF. 5).
FT CONFLICT 187 187 Q -> E (IN REF. 3 AND 5).
FT CONFLICT 201 201 N -> D (IN REF. 5).
FT CONFLICT 218 218 Q -> E (IN REF. 5).
FT CONFLICT 233 233 E -> Q (IN REF. 5).
FT CONFLICT 246 246 N -> D (IN REF. 5).
FT CONFLICT 256 256 E -> G (IN REF. 5).
```

```

FT  CONFLICT 260 260 N -> D (IN REF. 5).
FT  CONFLICT 266 266 N -> D (IN REF. 5).
FT  CONFLICT 280 280 Y -> W (IN REF. 5).
FT  CONFLICT 284 284 N -> S (IN REF. 5).
SQ  SEQUENCE 323 AA; 35404 MW; 69E8AA118D579A8B CRC64;

Query Match 64.5%; Score 923; DB 1; Length 323;
Best Local Similarity 69.8%; Pred. No. 8.4e-24;
Matches 169; Conservative 29; Mismatches 39; Indels 5; Gaps 2;

QY 28 VAEKLKEAFMDKT---HTC--PPCAPPELLGGPSVLFPPKPKDITLMISRTPEVTCVVD 82
DB 82 VAHPATNKVDKTVAPSTCSKPTCPPELLGGPSVFIFFPKPKDITLMISRTPEVTCVVD 141
QY 83 VSHEDPEKFNWVGVGEVHNATKPREQYNSTYRVSVTLTVLHODWLGKYEKCKVSN 142
DB 142 VSQDDPEVQFTWINNEQVTRAPPLREQOFNSTIRVSTPLTHODWLGKFEKCKVHN 201
QY 143 KALPAPIEKTISKAKQPREQVYTLPPSRDELTKNOVSLTCLVKGFYPSDIAVEWESNG 202
DB 202 KALPAPIEKTISKARGOPLKPVYTWGPPREELSSRSVSLTCMNGFYPSDISVEWENKNG 261
QY 203 QPENNYKTPPVLDSDGSGFLSKLVDSKRWQGNVFCVSMVHEALHNNHYTKQSLSLSP 262
DB 262 KAEDNYKTPPAVLDSGSGFLYLNKLSVPTSEWQSGDVFCTCSMVHEALHNNHYTKQSLSRSP 321
QY 263 GK 264
DB 322 GK 323

RESULT 6
GC2_CAVPO STANDARD; PRT; 329 AA.
AC P01862;
DT 21-JUL-1986 (Rel. 01, Created)
DT 21-JUL-1986 (Rel. 01, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE Ig gamma-2 chain C region.
OS Cavia porcellus (Guinea pig).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Hystricognathi; Caviidae; Cavia.
OX NCBI_TaxID=10141;
RN [1]
RP SEQUENCE OF 1-3.
RA Trischmann T.M.;
RL Submitted (APR-1975) to the PIR data bank.
RN [2]
RP SEQUENCE OF 4-68.
RX MEDLINE=71058471; PubMed=5538606;
RA Birshstein B.K., Hussain Q.Z., Cebra J.J.;
RT "Structure of heavy chain from strain 13 guinea pig
RT immunoglobulin-G(2). 3. Amino acid sequence of the region around the
RT half-cysteine joining heavy and light chains."
RL Biochemistry 10:18-25(1971).
RN [3]
RP SEQUENCE OF 69-133 AND 312-329.
RX MEDLINE=71058486; PubMed=5538616;
RA Turner K.J., Cebra J.J.;
RT "Structure of heavy chain from strain 13 guinea pig
RT immunoglobulin-G(2). II. Amino acid sequence of the carboxyl-terminal
RT and hinge region cyanogen bromide fragments."
RL Biochemistry 10:9-17(1971).
RN [4]
RP SEQUENCE OF 134-226.
RX MEDLINE=75036072; PubMed=4429665;
RA Tracey D.E., Cebra J.J.;
RT "Primary structure of the CH2 homology region from guinea pig IgG2
RT antibodies."
RL Biochemistry 13:4796-4803(1974).
RN [5]
RP SEQUENCE OF 227-311.
RX MEDLINE=75036073; PubMed=4609467;

```

```

RA Trischmann T.M., Cebra J.J.;
RT "Primary structure of the CH3 homology region from guinea pig IgG2
RT antibodies."
RL Biochemistry 13:4804-4811(1974).
RN [6]
RP DISULFIDE BONDS.
RX MEDLINE=71058474; PubMed=4922544;
RA Oliveira B., Lamm M.E.;
RT "Interchain disulfide bridges of guinea pig gamma-2-immunoglobulin."
RL Biochemistry 10:26-31(1971).
CC -I- MISCELLANEOUS: THIS CHAIN WAS ISOLATED FROM POOLED SERUM OF STRAIN
CC 13 INBRED GUINEA PIGS.
DR PIR; A02151; G2GP.
DR HSP; P01842; 7FAB.
DR InterPro; IPR003006; Ig_MHC.
DR InterPro; IPR003597; Ig-CL.
DR InterPro; IPR003600; Ig_Like.
DR Pfam; PF00047; Ig; 2.
DR SMART; SM00410; Ig_Like; 1.
DR SMART; SM00407; IGc1; 2.
DR PROSITE; PS00290; IG_MHC; 1.
KW Immunoglobulin domain; Immunoglobulin C region; Glycoprotein.
FT NON_TER 1 1
FT DISULFID 16 16 INTERCHAIN (WITH A LIGHT CHAIN).
FT DISULFID 28 79
FT DISULFID 105 105 INTERCHAIN (WITH A HEAVY CHAIN).
FT DISULFID 107 107 INTERCHAIN (WITH A HEAVY CHAIN).
FT DISULFID 110 110 INTERCHAIN (WITH A HEAVY CHAIN).
FT CARBOHYD 178 178 N-LINKED (GLCNAC. . .).
FT DISULFID 248 308
SQ SEQUENCE 329 AA; 36074 MW; 5D231B7164D1FBA9 CRC64;

Query Match 62.9%; Score 899; DB 1; Length 329;
Best Local Similarity 67.1%; Pred. No. 5.8e-23;
Matches 167; Conservative 25; Mismatches 44; Indels 13; Gaps 2;

QY 28 VAEKLKEAFMDKT-----HTCPCPAPELLGGPSVLFPPKPKDITLMISRTPEV 76
DB 81 VAHPASTTKVDKTVETPIPTZPBPCPCPPENLGGPSVFIFFPKPKDITLMISLTPRV 140
QY 77 TCVVVDVSHEDPEVKFNWYGVGEVHNATKPREQYNSTYRVSVTLTVLHODWLGKREY 136
DB 141 TCVVVDVSDQDEPEVQFTWFVDNKPVGNAETKPRVEQYNTTFRVESVLPFIQHODWLGKREF 200
QY 137 KCKVSNKALPAPIEKTISKAKQPREQVYTLPPSRDELTKNOVSLTCLVKGFYPSDIAY 196
DB 201 KCKVYNKALPAPIEKTISKAKGAPRPMDVYTLPPSRDELTSKSVSVTCLLIINFFPADIHV 260
QY 197 EWESNGQP--ENNYKTTTPVLDSDGSGFLYSLTVDKSRWQGNVFCVSMVHEALHNNHYT 254
DB 261 EWASNRVPVSEKYEKNTPTPIEDADGSGYFLYSLTVDKSAWDQGTVTYTCVSMVHEALHNNHYT 320
QY 255 QKSLSLSPG 263
DB 321 QKAISRSPG 329

RESULT 7
GC3_MOUSE
ID GC3_MOUSE STANDARD; PRT; 329 AA.
AC P22436;
DT 01-AUG-1991 (Rel. 19, Created)
DT 01-AUG-1991 (Rel. 19, Last sequence update)
DE 16-OCT-2001 (Rel. 40, Last annotation update)
DE Ig gamma-3 chain C region, secreted form.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=85027161; PubMed=6092053;

```

```
RA Wels J.A., Word C.J., Rimm D., Der-Balan G.P., Martinez H.M.,
RA Tucker P.W., Blattner F.R.;
RT "Structural analysis of the murine IgG3 constant region gene.";
RL EMOB J. 3:2041-2046(1984).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; J00451; -; NOT_ANNOTATED_CDS.
DR PIR; B02156; G3MSC.
DR HSP; P01857; 1FC1.
DR InterPro; IPR003006; Ig_MHC.
DR InterPro; IPR003597; Ig-cl.
DR InterPro; IPR003600; Ig_Like.
DR Pfam; PF00047; Ig; 3.
DR SMART; SM00410; IG_Like; 1.
DR SMART; SM00407; IGcl; 2.
DR PROSITE; PS00290; IG_MHC; 1.
KW Immunoglobulin domain; Immunoglobulin C region; Glycoprotein;
KW Transmembrane; Alternative splicing.
FT NON_TER 1 1
FT DOMAIN 1 97 CH1.
FT DOMAIN 98 113 HINGE.
FT DOMAIN 114 223 CH2.
FT DOMAIN 224 327 CH3.
SO SEQUENCE 329 AA; 36228 MW; F45827174182BAD6 CRC64;

Query Match 59.7%; Score 853; DB 1; Length 329;
Best Local Similarity 49.4%; Pred. No. 2.1e-21;
Matches 157; Conservative 41; Mismatches 48; Indels 72; Gaps 5;

QY 3 LKAFYDK-----VAEKLKEAF-----DWLKAFYDKVAEKLK 33
DB 28 VKGYPEPVTVMKNYGALSGVRVSSVLQSGFYSLSLVTVPSSTW----- 74
QY 34 EAFMDKTHTC-----PP-----CPAPELLGSPSVFLPPRPKD 66
DB 75 ---PSQTVICNVHPASKTELIKRIEIPRPSTPPGSSCPGNILGSPSVFIIPRPKD 131
QY 67 TLMISRTPEVTCVVDVSHEDPEVKFNMYVDGVEVHNAKTPREEQYNSTYRVSVLTNL 126
DB 132 ALMISLTPKVTVCVVDVEDDPDVHVSFVDNKEVHTAQTQPREAQYNSTFRVVSALPIQ 191
QY 127 HDWLNGKEYCKVSNKALPAPIETKISKAGQPREPOVYTLPPSRDELTKNQVSLTCLV 186
DB 192 HQDWMRGKEFKCKVNNKALPAPIETISKPRGATQTPQVYTIPTPPREOMSKKVSILTCLV 251
QY 187 KGFYPSDITAVESWGQENNYKTPPVLDSDGSFFLYSKLTVDKSRWQGNVFCSCVMH 246
DB 252 TNFSEATSVSEWERNLEQDKNTPPILDSDGTFFLYSKLTVDTSWLOGEIETCSVYH 311
QY 247 EALHNHYTQKSLSLSPGK 264
DB 312 EALHNHYTQKSLSLSPGK 329

RESULT 8
GCB_RAT
ID GCB_RAT STANDARD; PRT; 333 AA.
AC P20761;
DT 01-FEB-1991 (Rel. 17, Created)
DT 01-FEB-1991 (Rel. 17, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE Ig gamma-2B chain C region.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
```

```

[1]
SEQUENCE FROM N.A.
RP MEDLINE-89232738; Pubmed-3149946;
RX Brueggemann M.;
RA "Evolution of the rat immunoglobulin gamma heavy chain gene family.";
RL Gene 74:473-482(1988).
DR PIR; PS0018; PS0018.
DR HSP; P01842; 7FAB.
DR InterPro; IPR003006; Ig_MHC.
DR InterPro; IPR003597; Ig-cl.
DR InterPro; IPR003600; Ig_Like.
DR Pfam; PF00047; Ig; 3.
DR SMART; SM00410; IG_Like; 1.
DR SMART; SM00407; IGcl; 2.
DR PROSITE; PS00290; IG_MHC; 1.
KW Immunoglobulin domain; Immunoglobulin C region.
FT NON_TER 1 1
FT DISULFID 15 15 INTERCHAIN (WITH A LIGHT CHAIN).
FT DISULFID 27 80
FT DISULFID 106 106 INTERCHAIN (WITH A HEAVY CHAIN).
FT DISULFID 109 109 INTERCHAIN (WITH A HEAVY CHAIN).
FT DISULFID 112 112 INTERCHAIN (WITH A HEAVY CHAIN).
FT DISULFID 115 115 INTERCHAIN (WITH A HEAVY CHAIN).
FT DISULFID 147 207
FT DISULFID 253 311
SQ SEQUENCE 333 AA; 36497 MW; 55F8B64D48D460A6 CRC64;

Query Match 59.4%; Score 849; DB 1; Length 333;
Best Local Similarity 63.2%; Pred. No. 2.9e-21;
Matches 153; Conservative 34; Mismatches 52; Indels 3; Gaps 1;

QY 26 DKVAEKLKEAFMDKTHTC---PCPAPELLGSPSVFLPPRPKDTLMISRTPEVTCVVD 82
DB 92 DKVERRNGGIGCHKPTCTCHKCPVPELPGSPSVFIIPRPKDLLISQNAKVTGVVD 151
QY 83 VSHEDPEVKFNMYVDGVEVHNAKTPREEQYNSTYRVSVLTNLHDWLNGKEYCKVSN 142
DB 152 VSEEPDVQFSWFVNNVEVHTAQTQPREQYNSTFRVVSALPIQHDWMRGKEFKCKVNN 211
QY 143 KALPAPIETKISKAGQPREPOVYTLPPSRDELTKNQVSLTCLVKGFYPSDITAVESWG 202
DB 212 KALPSPIETKISKPRGLVRKPVYVMGPTTEQTQVTLTSGTSGFLPDIGVWTSNG 271
QY 203 QPENNYKTPPVLDSDGSFFLYSKLTVDKSRWQGNVFCSCVMHAEALHNHYTQKSLSLSP 262
DB 272 HIERNKTEPVMDSGSEFMYSKLVNVERSDRSRAPEVCSVWHEGLNHHVKSISRPP 331
QY 263 GK 264
DB 332 GK 333

RESULT 9
GCB_MOUSE
ID GCB_MOUSE STANDARD; PRT; 398 AA.
AC P03987;
DT 23-OCT-1986 (Rel. 02, Created)
DT 01-AUG-1991 (Rel. 19, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE Ig gamma-3 chain C region, membrane-bound form.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=85027161; Pubmed=6092053;
RA Wels J.A., Word C.J., Rimm D., Der-Balan G.P., Martinez H.M.,
RA Tucker P.W., Blattner F.R.;
RT "Structural analysis of the murine IgG3 constant region gene.";
RL EMOB J. 3:2041-2046(1984).
RN [2]
RP SEQUENCE OF 328-398 FROM N.A.
```

```

RX MEDLINE=84041483; PubMed=6314258;
RA Komaromy M., Clayton L., Rogers J., Robertson S., Kettman J.,
RA Wall R.;
RT "The structure of the mouse immunoglobulin in gamma 3 membrane gene
segment.";
RL Nucleic Acids Res. 11:6775-6785(1983).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; J00451; AAB59655.1;
DR EMBL; V01526; CAA24767.1; ALT_SEQ.
DR PIR; A02155; G3MSM.
DR HSSP; P01857; LFC1.
DR InterPro; IPR003006; Ig_MHC.
DR InterPro; IPR003597; Ig_cl.
DR InterPro; IPR003600; Ig_like.
DR Pfam; PF00047; Ig; 3.
DR SMART; SM00410; Ig_like; 1.
DR SMART; SM00407; IGcl; 2.
DR PROSITE; PS00290; IG_MHC; 1.
KW Immunoglobulin domain; Immunoglobulin C region; Glycoprotein;
KW Transmembrane; Alternative splicing.
FT NON_TER 1
FT DOMAIN 1 97 CH1.
FT DOMAIN 98 113 HINGE.
FT DOMAIN 114 223 CH2.
FT DOMAIN 224 327 CH3.
FT TRANSMEM 346 362
FT DOMAIN 363 398
FT CONFLICT 333 333
FT CONFLICT 342 342 E -> G (IN REF. 2).
FT CONFLICT 388 388 P -> F (IN REF. 2).
SQ SEQUENCE 398 AA; 43929 MW; CF7F264B50A41B95 CRC64;

Query Match 58.9%; Score 842; DB 1; Length 398;
Best Local Similarity 52.7%; Pred. No. 8.9e-21;
Matches 134; Conservative 37; Mismatches 46; Indels 55; Gaps 4;

QY 10 VAELKKEAF-----DWLKAFYDKVAEKLKAEAFMDKTHTC----- 43
DB 52 VSSVLQSGFSLSSLVTPSSW-----PSQIVICNVAHPASKTELIKR 95
QY 44 -----PP---CPAPELLGGPSVFLFPPKPKDTLMISRTPEVTCVVDVSHEDPEV 90
DB 96 IEPRIPKPTTPGSSCPPGNILGGPSVFIAPPKPKDAMLSLTPKVTVCVVDVSEDDPDV 155
QY 91 KFNWYDGVGVHNAKTPREEOYNSTYRVSVLTVLHODWLNKKEYCKVSNKALPAPTE 150
DB 156 HVSFWFVNKEVHTAWTPQREAQNSTFRVVSALPIQHDQMRGKCFKCVNKNALPAPTE 215
QY 151 KTISKAKGQPREQVYTLPPSRDELTKNQVSLTCLVKGFYPSDIAVEWESNGQPENNYKT 210
DB 216 RTISKPKGRAQTPQVYTIPTPPQMSKKKSLTCLVTFNFEAISVWEVRNEGEODYKN 275
QY 211 TTPVLDSGDSFFLYSKLTVDKSWOQGVNFCVSVHMEALHNHYTKSLSP 262
DB 276 TTPILSDGTIFYLSKLTVDTSWLGQEIFTCVSVHMEALHNHHTKNSRSP 327

RESULT 10
GCL_RAT STANDARD; PRT; 326 AA.
AC P20759;
DT 01-FEB-1991 (Rel. 17, Created)
DT 01-FEB-1991 (Rel. 17, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE Ig gamma-1 chain C region.
```

```

OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_Taxid=10116;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=89232738; PubMed=3149946;
RA Brueggemann M.;
RT "Evolution of the rat immunoglobulin gamma heavy-chain gene family.";
RL Gene 74:473-482(1988).
DR PIR; PS0017; PS0017.
DR HSSP; P01842; 7FAB.
DR InterPro; IPR003006; Ig_MHC.
DR InterPro; IPR003597; Ig_cl.
DR Pfam; PF00047; Ig; 3.
DR SMART; SM00407; IGcl; 2.
DR PROSITE; PS00290; IG_MHC; 1.
KW Immunoglobulin domain; Immunoglobulin C region; Glycoprotein.
FT NON_TER 1
FT DOMAIN 1 97 CH1.
FT DOMAIN 98 112 HINGE.
FT DOMAIN 113 219 CH2.
FT DOMAIN 220 326 CH3.
FT DISULFID 27 82
FT DISULFID 102 102 INTERCHAIN (WITH A HEAVY CHAIN).
FT DISULFID 106 106 INTERCHAIN (WITH A HEAVY CHAIN).
FT DISULFID 109 109 INTERCHAIN (WITH A HEAVY CHAIN).
FT DISULFID 111 111 INTERCHAIN (WITH A HEAVY CHAIN).
FT DISULFID 140 200
FT DISULFID 246 304
FT CARBOHYD 176 176
SQ SEQUENCE 326 AA; 35946 MW; 013BAB45EF49B9DA CRC64;

Query Match 57.8%; Score 826; DB 1; Length 326;
Best Local Similarity 43.8%; Pred. No. 1.7e-20;
Matches 145; Conservative 45; Mismatches 39; Indels 102; Gaps 3;

QY 3 LKAFY----- 7
DB 29 VKGYPEPVTVTNSGALSSGVHTPPAVLQSLGTLTSSVTVPSSWPSQVTCNVAPHA 88
QY 8 --DKVAELKAEAFDMLKAFYDKVAEKLKAEAFMDKTHTCPPAPELGG----- 54
DB 89 SSTKVDKKI-----VPRNCGDGCKPCICTGSE 115
QY 55 -PSVFLFPPKPKDTLMISRTPEVTCVVDVSHEDPEVKENWYDGVGVHNAKTPREEOY 113
DB 116 VSSVFIAPPKPKDVLITLTPKVTVCVVDISQDDPEVHESWFDVDEVTATQTRPPEQF 175
QY 114 NSTYRVSVSLTVLHODWLNKKEYCKVSNKALPAPTEKTIISKAKGQPREQVYTLPPSRD 173
DB 176 NSTFERSVSELPILHODWLNKTRFRCKVTSAAPSPLEKTIISKPEGRQVPHVYTHSPTE 235
QY 174 ELTKNQVSLTCLVKGFYPSDIAVEWESNGQPENNYKTTPPVLDSDGSFFLYSKLTVDKSR 233
DB 236 EMTQNEVSTICVMKGFYPPDIYVEWMQMQPOENYKNTPTPTMDTDCGYFLYSLKNVKKK 295
QY 234 WOQGVNFCVSVHMEALHNHYTKSLSPCK 264
DB 296 WOQGVNFCVSVHMEALHNHYTKSLSPCK 326

RESULT 11
GCL_MOUSE STANDARD; PRT; 324 AA.
AC P01868;
DT 21-JUL-1986 (Rel. 01, Created)
DT 21-JUL-1986 (Rel. 01, Last sequence update)
DT 30-MAY-2000 (Rel. 39, Last annotation update)
DE Ig gamma-1 chain C region.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
```



OX NCBI\_TaxID=10090;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=80045036; PubMed=115593;  
 RA Honjo T., Obata M., Yamawaki-Kataoka Y., Kataoka T., Kawakami T.,  
 RA Takahashi N., Mano Y.;  
 RT "Cloning and complete nucleotide sequence of mouse immunoglobulin  
 RT gamma 1 chain gene."  
 RL Cell 18:559-568(1979).  
 RN [2]  
 RP SEQUENCE OF 76-324 FROM N.A. (MYELOMA PROTEIN MOPC 31C).  
 RX MEDLINE=80202559; PubMed=6769752;  
 RA Obata M., Yamawaki-Kataoka Y., Takahashi N., Kataoka T., Shimizu A.,  
 RA Mano Y., Seidman J.G., Peterlin B.M., Leder P., Honjo T.;  
 RT "Immunoglobulin gamma 1 heavy chain gene: structural gene sequences  
 RT cloned in a bacterial plasmid."  
 RL Gene 9:87-97(1980).  
 RN [3]  
 RP SEQUENCE OF 70-322 FROM N.A. (MYELOMA PROTEIN MOPC 21).  
 RX MEDLINE=80012837; PubMed=113776;  
 RA Rogers J., Clarke P., Salser W.;  
 RT "Sequence analysis of cloned cDNA encoding part of an immunoglobulin  
 RT heavy chain."  
 RL Nucleic Acids Res. 6:3305-3321(1979).  
 RN [4]  
 RP SEQUENCE (MYELOMA PROTEIN MOPC 21).  
 RX MEDLINE=78242288; PubMed=98524;  
 RA Adetugbo K.;  
 RT "Evolution of immunoglobulin subclasses. Primary structure of a  
 RT murine myeloma gamma chain."  
 RL J. Biol. Chem. 253:6068-6075(1978).  
 RN [5]  
 RP DISULFIDE BONDS (MOPC 21).  
 RX MEDLINE=73008889; PubMed=5073237;  
 RA Svasti J., Milstein C.;  
 RT "The disulphide bridges of a mouse immunoglobulin G1 protein."  
 RL Biochem. J. 126:837-850(1972).  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announcement/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL; V00793; CAA24172.1; -  
 DR EMBL; V00793; CAA24173.1; -  
 DR EMBL; V00793; CAA24174.1; -  
 DR EMBL; V00793; CAA24175.1; -  
 DR EMBL; V00795; CAA24176.1; -  
 DR PIR; A02159; GLMS  
 DR HSP; P01842; 7FAB.  
 DR GlycoSuiteDB; P01868; -  
 DR MGD; MGI:96446; Igh-4.  
 DR InterPro; IPR003006; Ig\_MHC.  
 DR InterPro; IPR003597; Ig\_cl.  
 DR Pfam; PF00047; Ig; 3.  
 DR SMART; SM00407; Igcl; 2.  
 DR PROSITE; P500290; IG\_MHC; 1.  
 DR Immunoglobulin domain; Immunoglobulin C region; Glycoprotein;  
 KW Alternative splicing.  
 FT NON\_TER 1  
 FT DOMAIN 1 97 CH1.  
 FT DOMAIN 98 110 HINGE.  
 FT DOMAIN 111 217 CH2.  
 FT DOMAIN 218 324 CH3.  
 FT DISULFID 27 82  
 FT DISULFID 102 102 INTERCHAIN (WITH A LIGHT CHAIN).  
 FT DISULFID 104 104 INTERCHAIN (WITH A HEAVY CHAIN).  
 FT DISULFID 107 107 INTERCHAIN (WITH A HEAVY CHAIN).  
 FT DISULFID 109 109 INTERCHAIN (WITH A HEAVY CHAIN).  
 FT DISULFID 138 138

FT CARBOHYD 174 174 N-LINKED (GLCNAC. . .).  
 FT DISULFID 244 302 /FTIG-CAR\_000055.  
 FT MOD\_RES 324 324 REMOVED POST-TRANSLATIONALLY.  
 FT CONFLICT 276 276 N -> D (IN REF. 3).  
 FT CONFLICT 278 278 N -> D (IN REF. 3).  
 SQ SEQUENCE 324 AA; 35704 MW; A338812F3D1F2C93 CRC64;  
 Query Match 57.6%; Score 823; DB 1; Length 324;  
 Best Local Similarity 44.9%; Pred. No. 2e-20;  
 Matches 144; Conservative 51; Mismatches 42; Indels 84; Gaps 4;  
 QY 3 LKAFY-----  
 Db 29 VKGYFPEPTVTWNSGSLSSGVHTFPVAVLQSLDLYTLSSSVTPSPRPSETVTCNVAHPA 88  
 QY 8 --DKVAEKLKEAFDWLKAFLKAEAFMDKTHTCPP--CPAPELGLGSPSVFLFPKP 63  
 Db 89 SSTKVDKKI-----VPRDGGCKPCICTVPEV---SSVFIFPPK 123  
 QY 64 PKDTLMISRTPEVTCVVVDVSHEDPEVKFNWYVDGVEVHNAKTKPREEQYNSTYRVVSVL 123  
 Db 124 PRDVLTIITLTPKVCVVVDISKDDPEVQSFVDDVEVHTAQTQPREQFNSTFRSVSEL 183  
 QY 124 TVLHDLNKGKEYCKVSNKALPAPIEKTISKAKGQPREPQVYTLPPSRDELTKNQVSLT 183  
 Db 184 PIMHQDLNKGKEYCKVSNKALPAPIEKTISKAKGQPREPQVYTLPPSRDELTKNQVSLT 243  
 QY 184 CLVKGFPYSDIAEVESNGQPENNYKTTPPVLDSDGSFFLYSKLTVDKSRWQQGNVPSCS 243  
 Db 244 CMITDFPEDITVEMQWNGQPAENYKNTQPIIMNTNGSYFYISKLNVQKSNWEAGNTFTCS 303  
 QY 244 VMHEALHNHYTOKSLSLSPGK 264  
 Db 304 VLHEGLNHHTEKLSHSPGK 324  
 RESULT 12  
 GCC\_RAT  
 ID GCC\_RAT STANDARD; PRT; 329 AA.  
 AC P20762;  
 DT 01-FEB-1991 (Rel. 17, Created)  
 DT 01-FEB-1991 (Rel. 17, Last sequence update)  
 DT 15-JUL-1999 (Rel. 38, Last annotation update)  
 DE Ig gamma-2C chain C region.  
 OS Rattus norvegicus (Rat).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.  
 OX NCBI\_TaxID=10116;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=88166903; PubMed=3127222;  
 RA Brueggemann M., Delmastro-Galfre P., Waldmann H., Calabi F.;  
 RT "Sequence of a rat immunoglobulin gamma 2c heavy chain constant  
 RT region cDNA: extensive homology to mouse gamma 3."  
 RL Eur. J. Immunol. 18:317-319(1988).  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announcement/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL; X07189; CAA30169.1; -  
 DR PIR; S00847; S00847.  
 DR HSP; P01842; 7FAB.  
 DR InterPro; IPR003006; Ig\_MHC.  
 DR InterPro; IPR003597; Ig-cl.  
 DR InterPro; IPR003600; Ig\_like.  
 DR Pfam; PF00047; Ig; 2.  
 DR SMART; SM00410; Ig\_like; 1.

```
DR SMART; SM00407; IGC1; 2;
KW PROSITE; PS00290; IG_MHC; 1;
KW Immunoglobulin domain; Immunoglobulin C region.
FT NON_TER 1 97
FT DOMAIN 1 97 CHI.
FT DOMAIN 98 113 HINGE.
FT DOMAIN 114 222 CHI.
FT DOMAIN 223 329 CHI.
FT DISULFID 15 15 INTERCHAIN (WITH A LIGHT CHAIN).
FT DISULFID 27 82 INTERCHAIN (WITH A LIGHT CHAIN).
FT DISULFID 111 111 INTERCHAIN (WITH A HEAVY CHAIN).
FT DISULFID 113 113 INTERCHAIN (WITH A HEAVY CHAIN).
FT DISULFID 143 203 INTERCHAIN (WITH A HEAVY CHAIN).
FT DISULFID 249 307
SQ SEQUENCE 329 AA; 36571 MW; 5FCD7B7933850773 CRC64;

Query Match 57.2%; Score 818; DB 1; Length 329;
Best Local Similarity 57.9%; Pred. No. 3.2e-20;
Matches 146; Conservative 42; Mismatches 38; Indels 26; Gaps 2;

QY 39 KTHTC-----PP---CPAPELLGGPSVFLPDKPKDTLMISR 72
DB 78 QTVCSVAHPATKSNLRIKRIEPRRPKPRPTDSCDNLGRSPVFPKPKDILMIL 137
QY 73 TPEVTGVVVDVSHDEPKENYVDGVEVHNAKTPREEQYNSTYRVSVLTVLHQDWLN 132
DB 138 TPKVTGVVVDVSEEDPVQFSWFVDNRVFTAQTPHEEQNLGTRFVSTLHIHQDWS 197
QY 133 GKEYCKVSNKALPAPIEKTISKAGQPREQVITLPSRDELTKNQVSLCLYKGFYPS 192
DB 198 GKEFKCKVNNKDLSPIEKTISKGRKARTPOVYITPPPREQMSKNKVSLSCTWYTFYA 257
QY 193 DIAVEWESNGPENNKYKTPPVLDSGDSFFLYSKLTVDKSRWQGNVFCSCVMHEALHNH 252
DB 258 SISVEWERNGELEQDYKNTLPVLDSDESIFYLSKLSVDTDSMRGDIYTCVWHEALHNH 317
QY 253 YTKSLSLSPGK 264
DB 318 HTQKNLSRSPGK 329

RESULT 13
GCIM_MOUSE STANDARD; PRT; 393 AA.
AC P01869;
DT 21-JUL-1986 (Rel. 01, Created)
DT 01-AUG-1991 (Rel. 19, Last sequence update)
DT 30-MAY-2000 (Rel. 39, Last annotation update)
DE I9 gamma-1 chain C region, membrane-bound form.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=80045036; PubMed=115593;
RA Honjo T., Obata M., Yamawaki-Kataoka Y., Kataoka T., Kawakami T.,
RA Takahashi N., Mano Y.;
RT "Cloning and complete nucleotide sequence of mouse immunoglobulin
RT gamma 1 chain gene.";
RL Cell 18:559-568(1979).
[2]
RP SEQUENCE OF 323-393 FROM N.A.
RX MEDLINE=82197626; PubMed=6804950;
RA Tyler B.M., Cowman A.F., Gerondakis S.D., Adams J.M., Bernard O.;
RT "mRNA for surface immunoglobulin gamma chains encodes a highly
RT conserved transmembrane sequence and a 28-residue intracellular
RT domain.";
RL Proc. Natl. Acad. Sci. U.S.A. 79:2008-2012(1982).
[3]
RP SEQUENCE OF 323-366 FROM N.A.
RX MEDLINE=82115295; PubMed=6799207;
RA Rogers J., Choi E., Souza L., Carter C., Word C.J., Kuehl M.,

Eisenberg D., Wall R.;
"Gene segments encoding transmembrane carboxyl termini of
immunoglobulin gamma chains.";
Cell 26:19-27(1981).
[4]
RP SEQUENCE OF 1-44 FROM N.A.
RX MEDLINE=82222190; PubMed=6283537;
RA Yamawaki-Kataoka Y., Nakai S., Miyata T., Honjo T.;
RT "Nucleotide sequences of gene segments encoding membrane domains of
immunoglobulin gamma chains.";
Proc. Natl. Acad. Sci. U.S.A. 79:2623-2627(1982).
-1- ALTERNATIVE PRODUCTS: CELL LINES PRODUCING IGG CONTAIN TWO MRNA
SPECIES FOR IG GAMMA CHAINS. THE MAJOR SPECIES ENCODES SECRETED-
GAMMA CHAINS. A LESS ABUNDANT SPECIES APPEARS TO ENCODE MEMBRANE-
BOUND CHAINS IN THAT IT CONTAINS AN ALTERNATIVE 3' END, ENCODED
IN SEPARATE EXONS, THAT IS HOMOLOGOUS WITH THE MEMBRANE-BOUND
SEGMENT OF MU CHAINS.
-----
THIS SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch).
-----
EMBL; V00793; CAA24172.1; -
EMBL; V00793; CAA24173.1; -
EMBL; V00793; CAA24174.1; -
PIR; B02159; GIMSM.
HSP; P01842; 7FAB.
MGD; MGI:96446; Igh-4.
InterPro; IPR003006; Iq_MHC.
InterPro; IPR003597; Iq-cl.
Pfam; PF00047; Iq; 3.
SMART; SM00407; IGC1; 2.
PROSITE; PS00290; IG_MHC; 1.
Immunoglobulin domain; Immunoglobulin C region; Glycoprotein;
Alternative splicing; Transmembrane.
NON_TER 1 97
DOMAIN 1 97 CHI.
DOMAIN 98 110 HINGE.
DOMAIN 111 217 CHI.
DOMAIN 218 324 CHI.
DISULFID 27 82
DISULFID 102 102 INTERCHAIN (WITH A LIGHT CHAIN).
DISULFID 104 104 INTERCHAIN (WITH A HEAVY CHAIN).
DISULFID 107 107 INTERCHAIN (WITH A HEAVY CHAIN).
DISULFID 109 109 INTERCHAIN (WITH A HEAVY CHAIN).
DISULFID 138 198
CARBOHYD 174 174 N-LINKED (GLCNAC. . .).
DISULFID 244 302 POTENTIAL.
TRANSMEM 340 357 CYTOPLASMIC (POTENTIAL).
DOMAIN 358 393
SEQUENCE 393 AA; 43386 MW; 4CC88343B7AICE27 CRC64;

Query Match 57.2%; Score 818; DB 1; Length 393;
Best Local Similarity 59.6%; Pred. No. 5.5e-20;
Matches 140; Conservative 48; Mismatches 42; Indels 5; Gaps 2;

QY 31 KLKEAFMDKTHTCPP--CPAPELLGGPSVFLPDKPKDTLMISRPEVTCVVVDVSHEDP 88
DB 92 KVDKKIVPRDCGCKPCICTPEV---SSVFFPPKPKDVLTLTPKVTVCVVVDISKDDP 148
QY 89 EVKENWYVDGVEVINAKTPREEQYNSTYRVSVLTVLHQDWLNGKEYCKVSNKALPAP 148
DB 149 EVQFSWFVDDVEVHTAQTQPREQFNSTFRSVSELPIMHQDWLNGKEYCKVSNKALPAP 208
QY 149 IEKTISKAGQPREQVITLPSRDELTKNQVSLTCLVKGFYPSDIAVEWESNGOPENNY 208
DB 209 IEKTISKGRKAPQVYITPPPREQMAKDKVSLTCTITDEFFEDITVEWQWNGQPAENY 268
QY 209 KTTPPVLDSGDSFFLYSKLTVDKSRWQGNVFCSCVMHEALHNHHTYKSLSPG 263
```







```

Db 242 KSCDKTHTCPCPAPBELLGGPSVFLPPPKPKDITLMISRTPEVTCVVVDVSHEDPEVKFNW 301
QY 95 YVDGVEVHNAKTKPREQYNSTYRVVSVLTVLHQDLNMGKEYCKVSNKALPAPIEKTIS 154
Db 302 YVDGVEVHNAKTKPREQYNSTYRVVSVLTVLHQDLNMGKEYCKVSNKALPAPIEKTIS 361
QY 155 KAKGQPREPQVYTLPPSDELTKNQVSLTCLVKGFYPSDIAVEWESNGOPENNYKTTTPV 214
Db 362 KAKGQPREPQVYTLPPSDELTKNQVSLTCLVKGFYPSDIAVEWESNGOPENNYKTTTPV 421
QY 215 LDSGGSFFLYSKLTVDKSRWQGNVFCSCVMHEALHNHYTQKSLSLSPGK 264
Db 422 LDSGGSFFLYSKLTVDKSRWQGNVFCSCVMHEALHNHYTQKSLSLSPGK 471

RESULT 2
Q96P08 PRELIMINARY; PRT; 701 AA.
AC Q96P08;
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-MAR-2002 (TrEMBLrel. 20, Last annotation update)
DE Factor VII active site mutant immunofugate.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=21477448; PubMed=11593034;
RA Hu Z., Garen A.;
RT "Targeting tissue factor on tumor vascular endothelial cells and tumor
RT cells for immunotherapy in mouse models of prostatic cancer.";
RT Proc. Natl. Acad. Sci. U.S.A. 98:12180-12185(2001).
DR EMBL: AF272774; AAK58686.1; -
DR InterPro: IPR000152; Asx_Hydroxyl.
DR InterPro: IPR000561; EGF-like.
DR InterPro: IPR000742; EGF-2.
DR InterPro: IPR001881; EGF_Ca.
DR InterPro: IPR003006; Ig_MHC.
DR InterPro: IPR001254; Ser_protease_Try.
DR InterPro: IPR000294; VitK_dep_GUA.
DR Pfam: PF00008; EGF; 2.
DR Pfam: PF00594; gla; 1.
DR Pfam: PF00047; ig; 2.
DR Pfam: PF00089; trypsin; 1.
DR SMART: SM00181; EGF; 2.
DR PROSITE: PS00010; ASX_HYDROXYL; UNKNOWN_1.
DR PROSITE: PS00022; EGF_1; UNKNOWN_1.
DR PROSITE: PS01186; EGF_2; UNKNOWN_1.
DR PROSITE: PS01187; EGF_CA; UNKNOWN_1.
DR PROSITE: PS00011; GLU_CARBOXYLATION; UNKNOWN_1.
DR PROSITE: PS00290; IG_MHC; UNKNOWN_1.
DR PROSITE: PS0240; TRYPSIN_DOM; 1.
DR PROSITE: PS00134; TRYPSIN_HIS; UNKNOWN_1.
DR PROSITE: PS00135; TRYPSIN_SER; UNKNOWN_1.
KW Hydrolase; Serine protease.
SQ SEQUENCE 701 AA; 77826 MW; 94AC6CEB42CC992F CRC64;

Query Match 87.5%; Score 1251; DB 4; Length 701;
Best Local Similarity 51.2%; Pred. No. 7.3e-41;
Matches 233; Conservative 8; Mismatches 9; Indels 205; Gaps 3;

QY 7 YDKVAE----- 12
Db 255 FDKIKNRNLIAVLGEHDLSEHDGDSRRVAQVIIPSTVPGTTNHDIALRLHQPVWL 314
QY 13 ----- 12
Db 315 TDHVVPLCLPERTSERTLAFVRESLVSGWQLDGRGATALELMVLNVPRLMTQDCLQQS 374
QY 13 ----- 12

```

```

Db 375 RKVGDSPNITEYMFAGYSKSGSKDCAGSGSPHATHYRTGWTYLTGIVSWGQCATVGHF 434
QY 13 ----KLEAFDWLKAFTYDKVAEKLKEA-----FMDKTHTCPCPAP 49
Db 435 GYTRVSGYIEWL-----QKLMRSEPRGVLLRAPFGSAEPKSCDKTHTTCPCPAP 486
QY 50 ELLGGPSVFLPPPKPKDITLMISRTPEVTCVVVDVSHEDPEVKFNWYVDGVEVHNAKTKPR 109
Db 487 ELLGGPSVFLPPPKPKDITLMISRTPEVTCVVVDVSHEDPEVKFNWYVDGVEVHNAKTKPR 546
QY 110 EEOYNSTYRVVSVLTVLHQDLNMGKEYCKVSNKALPAPIEKTISKAKGQPREPQVYTL 169
Db 547 EEOYNSTYRVVSVLTVLHQDLNMGKEYCKVSNKALPAPIEKTISKAKGQPREPQVYTL 606
QY 170 PSDELTKNQVSLTCLVKGFYPSDIAVEWESNGOPENNYKTTTPVLDSDGGSFFLYSKLTV 229
Db 607 PSDELTKNQVSLTCLVKGFYPSDIAVEWESNGOPENNYKTTTPVLDSDGGSFFLYSKLTV 666
QY 230 DKSRWQGNVFCSCVMHEALHNHYTQKSLSLSPGK 264
Db 667 DKSRWQGNVFCSCVMHEALHNHYTQKSLSLSPGK 701

RESULT 3
Q8TC63 PRELIMINARY; PRT; 473 AA.
AC Q8TC63;
DT 01-JUN-2002 (TrEMBLrel. 21, Created)
DT 01-JUN-2002 (TrEMBLrel. 21, Last sequence update)
DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE Hypothetical 52.0 kDa protein.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Srausberg R.;
RL Submitted (MAR-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL: BC025985; AAH25985.1; -
KW Hypothetical protein.
SQ SEQUENCE 473 AA; 51986 MW; E29920B09BA369F5 CRC64;

Query Match 80.8%; Score 1155; DB 4; Length 473;
Best Local Similarity 48.8%; Pred. No. 1.5e-37;
Matches 219; Conservative 22; Mismatches 22; Indels 188; Gaps 5;

QY 2 WLKA-----FYDKVAEKLKEAFDWL----- 21
Db 23 WVLRLQLQESGPGCLKKPSVTLSTCTVSGDSVASSSYNGWVRQPPGKLEWIGTINFS 82
QY 22 -----KAFYDKVAE----- 30
Db 83 GNMYYSPSLRSRVMTSADMSSENFYKLDSVTAAADTAVYYCAAGHLVMGFGAHWGQKLV 142
QY 31 -----KLKEAFDWL----- 38
Db 143 SVSPASTKGSPVPLAPCSRSTSESTAALGCLVKDYFPEPVTVSWNSGALTSGVHTTTPAV 202
QY 39 -----KTHT-----GPPCPAPBELLG 53
Db 203 LQSSGLYSLSSVTVTPSSSLGTCTYTCNVDPKPSNTKVKDKRVESKYGPPCSCPAPEFLG 262
QY 54 GPSVFLPPPKPKDITLMISRTPEVTCVVVDVSHEDPEVKFNWYVDGVEVHNAKTKPREQY 113
Db 263 GPSVFLPPPKPKDITLMISRTPEVTCVVVDVSHEDPEVKFNWYVDGVEVHNAKTKPREQY 322
QY 114 NSTYRVVSVLTVLHQDLNMGKEYCKVSNKALPAPIEKTISKAKGQPREPQVYTLPPSRD 173
Db 323 NSTYRVVSVLTVLHQDLNMGKEYCKVSNKALPAPIEKTISKAKGQPREPQVYTLPPSQE 382
QY 174 ELTKNQVSLTCLVKGFYPSDIAVEWESNGOPENNYKTTTPVLDSDGGSFFLYSKLTVDKSR 233

```

```

DE 383 EMKQVSLCLYKGFYPDIWESNGOPENNYKTPPVLDSDGSFFLYSLRTVDKSR 442
QY 234 WQGNVSCSVMHEALHNHYTKSLSPCK 264
DB 443 WOENVFSCSVMHEALHNHYTKSLSLGK 473

RESULT 4
Q95M34 PRELIMINARY; PRT; 337 AA.
AC Q95M34;
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE Immunoglobulin gamma 1 heavy chain constant region
DE (Fragment).
GN IGHC1.
OS Equus caballus (Horse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Perissodactyla; Equidae; Equus.
OX NCBI_TaxID=9796;
RN [1]
RP SEQUENCE FROM N.A.
RA Wagner B.;
RL Submitted (DEC-2000) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=98383416; PubMed=9717671;
RA Wagner B., Overesch G., Sheoran A., Holmes M., Richards C.,
RA Leibold W., Radbruch A.;
RT "Organization of the equine immunoglobulin heavy chain constant region
RT genes. III. Alignment of c-mu, c-gamma, c-epsilon and c-alpha genes.";
RL Immunobiology 199:105-119(1998).
DR EMBL; AJ300675; CAC4624.1; -.
DR InterPro; IPR003006; Ig_MHC.
DR Pfam; PF00047; Ig; 2.
DR PROSITE; PS00290; IG_MHC; UNKNOWN_2.
FT NON_TER
FT 1
SQ SEQUENCE 337 AA; 37438 MW; A60BF2B01DEFD1F6 CRC64;

Query Match 64.4%; Score 921; DB 6; Length 337;
Best Local Similarity 60.8%; Pred. No. 1.4e-28;
Matches 169; Conservative 36; Mismatches 42; Indels 31; Gaps 4;

QY 2 WLKAFY-----DKVAEKLKEAFDNLKAFYDKVKAELKEAFMDKTHTCPCPA 48
DB 76 WTSETYICNVVHAASNFVKDKRIEPIPN-----HOKVCDMSK-----CPKCPA 119
QY 49 PELLGGPSVFLFPKPKDLMISRTPEVTCVVVDVSHEDPEVKFNWYVDGVEVHNATKP 108
DB 120 PELLGGPSVFLFPNPKDLMITPTVTCVVVDVSDQENPDVKNWYMDGVEVETATRP 179
QY 109 REEQNSTYRVSVLTVLHODWLNKGEYKCKVSNKALPAPIEKTISKAKGQPREPVYTL 168
DB 180 KEEQFNSTYRVSVLRIQHODWLNKGEYKCKVSNQALPQPIERTITTKGRSQEPQVVL 239
QY 169 PPSRDELTKNOVSLTCLVKGFPSPDIWESNGQP--ENNYKTPPVLDSDGSFFLYSK 226
DB 240 APHPDELKSKSVTCLVKDIPPEINIEWQSQNPQLETKYSTTQAQDSDGSYFLYSK 299
QY 227 LTVDKSRWQGNVSCSVMHEALHNHYTKSLSPCK 264
DB 300 LSVDRNRWQGNVSCSVMHEALHNHYTKSNPKG 337

RESULT 5
Q99LC4 PRELIMINARY; PRT; 463 AA.
AC Q99LC4;
DT 01-JUN-2001 (TrEMBLrel. 17, Created)
DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)

```

```

DE Similar to RIKEN CDNA 1810060009 gene.
GN IGH-4.
OC Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RA Strausberg R.;
RL Submitted (FEB-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC003435; AAH03435.1; -.
DR HSSP; P01842; 7FAB.
DR MGD; MGI:96446; Igh-4.
DR InterPro; IPR003599; Ig.
DR InterPro; IPR003597; Ig_cl.
DR InterPro; IPR003600; Ig_like.
DR InterPro; IPR003006; Ig_MHC.
DR InterPro; IPR003596; Ig_v.
DR Pfam; PF00047; Ig; 4.
DR SMART; SM00409; IG; 2.
DR SMART; SM00407; IGcl; 2.
DR SMART; SM00406; IGV; 1.
DR SMART; SM00410; IG_like; 1.
DR PROSITE; PS00290; IG_MHC; UNKNOWN_1.
SQ SEQUENCE 463 AA; 51007 MW; EAA674C6BBC30783 CRC64;

Query Match 59.7%; Score 854; DB 11; Length 463;
Best Local Similarity 37.0%; Pred. No. 2.1e-25;
Matches 155; Conservative 50; Mismatches 48; Indels 166; Gaps 8;

QY 2 WLKAFYDKVAEKLKEAFDNLKAFY-----DKVAEKLK-----EAFMD----- 38
DB 55 WVK-----QRTQGLEWGEIYPGSGNTYSEKFKGKATLTDDKSSSTAYMHLSSITS 107
QY 39 -----HT-----KT----- 40
DB 108 EDSAVFYFCARSSYYSDLFAYWGQGLTVTSAKTPPSVYPLAPGSAQAQNSMVTGLCL 167
QY 41 -----HT----- 42
DB 168 VKGFPEPVTVWNSGSLSSGVHTFPAVLQSDLYLTLSSTVTPSPSTWPEVTCNVARPA 227
QY 43 -----CPP--CPAPELLGGPSVFLFPKPKDLMISRTPEVTCVVVDVSH 85
DB 228 SSTRVDKKIIPRDCGCKPCICTVPEV---SSVFIFPPKPKDVLITLTPKTCVVVDISK 284
QY 86 EDPEKFNWYVDGVEVHNATKPREEQYNSTYRVSVLTVLHODWLNKGEYKCKVSNKAL 145
DB 285 DDPEVQFSWFVDVETVTAQTPREQFNSTFRSVSELPIMHODWLNKGEYKCKVNSAAF 344
QY 146 PAPIEKTISKAKGQPREPVYTLPPSRDELTKNOVSLTCLVKGFPSPDIWESNGOPE 205
DB 345 PAPIEKTISKAKGQPREPVYTLPPSRDELTKNOVSLTCLVKGFPSPDIWESNGOPE 404
QY 206 NNYKTPPVLDSDGSFFLYSKLTVDKSRWQGNVSCSVMHEALHNHYTKSLSPCK 264
DB 405 ENYKTPQINDTSGSYFLYSLKLVKSNWAGNFTCSVLHDEGLHNHHHTSLSPCK 463

RESULT 6
Q8R3V9 PRELIMINARY; PRT; 469 AA.
AC Q8R3V9;
DT 01-JUN-2002 (TrEMBLrel. 21, Created)
DT 01-JUN-2002 (TrEMBLrel. 21, Last sequence update)
DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE Hypothetical 52.0 kDa protein.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.

```

DR MGD:



```
DR InterPro: IPR003599; Ig.
DR SMART: SM00409; IG; 2.
DR SMART: SM00407; IGcl; 3.
DR SMART: SM00406; IGv; 1.
DR SMART: SM00410; IG_Like; 1.
DR SMART: SM00407; IG; 2.
DR SMART: SM00409; IGcl; 3.
DR SMART: SM00406; IGv; 1.
DR SMART: SM00410; IG_Like; 1.
DR PROSITE: PS00290; IG_MHC; UNKNOWN_1.
SQ SEQUENCE 473 AA; 51699 MW; 9DED57A514475FBB CRC64;

Query Match      58.7%; Score 840; DB 11; Length 473;
Best Local Similarity 36.4%; Pred. No. 8.3e-25;
Matches 157; Conservative 46; Mismatches 50; Indels 178; Gaps 7;

QY 1 DWLK-----AFYDKVAEKLKEA----- 17
DB 54 MWKQRPQGLEWIGKIGPGSGSTYNN---EKFKGKATLTADKSSSTAYMQLSSTSEDS 110
QY 18 -----FDW----- 24
DB 111 AVIFCARSGYDIDWFAYWGGTTLTVSAKTTAPSVYPLAPVCGGTGSSVTGLGLVKGY 170
QY 25 Y----- 25
DB 171 FPEPVTLTWNSGSLSGVHTFPALLOGLYTLSSSVTVTSNTWPSQTIITCNVAHPASSTK 230
QY 26 -DKVAE-----KLKEAFMDKTHTCPAPPELLGGPSVFLFPPKPKDTLMISRT 73
DB 231 VDKKIEPRVITONPCPPLE-----CPPCAAPDLLGGPSVFIIPPRIKIDVLMISLS 282
QY 74 PEYTCVVVDYSHEDPEVKFNWYVDGVEVHNAKTKPREEQYNSTYRVVSVLTVTLQHDWNLG 133
DB 283 PMTCVVVDVSDDDPQVLSFVNVEVHTAQTQTHREDYNSLTRVVSALPIQHODWMSG 342
QY 134 KEYKCKVSKALPAPIETKISKARGOPREPQVYTLPPSRDELTKNOVSLTCLVKGFYPSD 193
DB 343 KEFKCKVNNRALPSPIETKISKRGVPVRAPQVYVLPPEAEETKKEFSLTCMITGLFPAE 402
QY 194 IAVENESNGQPNNTKTPPVLDSDGSFFLYSKLTVDKSRWQQGNVFCSMVEALHNNHY 253
DB 403 IAVDWTSGNTEQNKNTATVLDSDGSYFMYSKLRVQKSTWGRGSLFACSWHEGLHNL 462
QY 254 TQKSLSLSPCK 264
DB 463 TTKTISRSLGK 473

RESULT 9
Q99L25 PRELIMINARY; PRT; 473 AA.
AC Q99L25;
DT 01-JUN-2001 (TrEMBLrel. 17, Created)
DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE Similar to RIKEN CDNA 1810060009 gene.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RA Strausberg R.;
RL Submitted (FEB-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC003888; AA03888.1;
DR HSSP; P01842; 7FAB.
DR InterPro: IPR003599; Ig.
DR InterPro: IPR003597; Ig_cl.
DR InterPro: IPR003600; Ig_Like.
DR InterPro: IPR003006; Ig_MHC.
DR InterPro: IPR003596; Ig_v.

QY 254 TQKSLSLSPCK 264
DB 463 TTKTISRSLGK 473

RESULT 10
Q99L31 PRELIMINARY; PRT; 468 AA.
AC Q99L31;
DT 01-JUN-2001 (TrEMBLrel. 17, Created)
DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE Similar to RIKEN CDNA 1810060009 gene.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RA Strausberg R.;
RL Submitted (FEB-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC003878; AA03878.1;
DR HSSP; P01842; 7FAB.
DR InterPro: IPR003599; Ig.
DR InterPro: IPR003597; Ig_cl.
DR InterPro: IPR003600; Ig_Like.
DR InterPro: IPR003006; Ig_MHC.
DR InterPro: IPR003596; Ig_v.
DR Pfam; PF00047; Ig; 3.
DR SMART; SM00409; IG; 2.
DR SMART; SM00407; IGcl; 3.
DR SMART; SM00406; IGv; 1.
DR SMART; SM00410; IG_Like; 1.
```

```
DR Pfam; PF00047; Ig; 3.
DR SMART; SM00409; IG; 2.
DR SMART; SM00407; IGcl; 3.
DR SMART; SM00406; IGv; 1.
DR SMART; SM00410; IG_Like; 1.
DR PROSITE: PS00290; IG_MHC; UNKNOWN_1.
SQ SEQUENCE 473 AA; 52449 MW; BE9889B7986DA155 CRC64;

Query Match      58.5%; Score 836; DB 11; Length 473;
Best Local Similarity 37.6%; Pred. No. 1.2e-24;
Matches 160; Conservative 38; Mismatches 58; Indels 170; Gaps 7;

QY 2 WLKAFYDKVAEKLKEAFDWLKAFY-----DKVAEKLK-----BAFM----- 37
DB 55 WVK-----QRPEQGLEWIGYIYPRDGSSTKYNEKFKGKATLTADKSSSTAYMQLSLS 107
QY 38 ----- 40
DB 108 EDSAVCFCSRGSYIYGYLYFYDWGQGTITVSSAKTTAPSVYPLAPVCGDTTGSST 167
QY 41 -----HT----- 42
DB 168 LGCLVKGYFPEPVTLTWNSGSLSSGVHTFPAVLQSDLYTLSSSVTVTSWPSQITCNV 227
QY 43 -----CPP-----CPAPELLGGPSVFLFPPKPKDTLMISRTPEVTC 78
DB 228 AHPASSTKVDKKTIEPRGPTIKPCPCPCAPNLLGGPSVFIIPPRIKIDVLMISLSPWTC 287
QY 79 VYVDVSHEDPEVKFNWYVDGVEVHNAKTKPREEQYNSTYRVVSVLTVTLQHDWNLGKYEKC 138
DB 288 VYVDVSEDDPQVLSFVNVEVHTAQTQTHREDYNSLTRVVSALPIQHODWMSGKEFKC 347
QY 139 KVSNNKALPAPIETKISKARGOPREPQVYTLPPSRDELTKNOVSLTCLVKGFYPSDIAVEW 198
DB 348 KVNNKALPAPIETKISKRGVPVRAPQVYVLPPEAEETKKEFSLTCMITGLFPAE 407
QY 199 ESNQGNPNNTKTPPVLDSDGSFFLYSKLTVDKSRWQQGNVFCSMVEALHNNHYTKSL 258
DB 408 TNGKTELNYKNTPEVLDSDGSYFMYSKLRVQKSTWGRGSLFACSWHEGLHNNHTKSF 467
QY 259 SLSPGK 264
DB 468 SRTPGK 473

RESULT 10
Q99L31 PRELIMINARY; PRT; 468 AA.
AC Q99L31;
DT 01-JUN-2001 (TrEMBLrel. 17, Created)
DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE Similar to RIKEN CDNA 1810060009 gene.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RA Strausberg R.;
RL Submitted (FEB-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC003878; AA03878.1;
DR HSSP; P01842; 7FAB.
DR InterPro: IPR003599; Ig.
DR InterPro: IPR003597; Ig_cl.
DR InterPro: IPR003600; Ig_Like.
DR InterPro: IPR003006; Ig_MHC.
DR InterPro: IPR003596; Ig_v.
DR Pfam; PF00047; Ig; 3.
DR SMART; SM00409; IG; 2.
DR SMART; SM00407; IGcl; 3.
DR SMART; SM00406; IGv; 1.
DR SMART; SM00410; IG_Like; 1.
```

```
DR PROSITE; PS00290; IG_MHC; UNKNOWN_1.
SQ SEQUENCE 468 AA; 51661 MW; 96352328B3332ADB CRC64;

Query Match
Best Local Similarity 58.0%; Score 830; DB 11; Length 468;
Matches 158; Conservative 38; Mismatches 55; Indels 175; Gaps 7;

QY 2 WLKAFYDKVAEKLKEAFDMLKAFYD-----KVAEKLKEAFMDK----- 39
Db 55 WK-----QRPEQGLEWI-GWIDPEDGETKYAPK-----FQDKATITADTSSNTAYLQL 102
QY 40 -----THT----- 39
Db 103 SSLTSEDTAIYCARNLGYGYDYGQGTITVSSAKTTAPSVYPLAPVCGDTTGSSTV 162
QY 40 -----THT----- 42
Db 163 LGCLVKGYPEPVTLTWNSGSLSSGVHTFPALQSLGYTMSSSVTVPSSTWPSQITCNV 222
QY 43 -----CPP--CPAPELLGGPSVFLPPPKPKDTLMISRTPEVTC 78
Db 223 AHPASSTKVDDKIEPRGPTIKPCPKCKAPNLEGSPSVIFPPNKKDVLMLSLSPWYTC 282
QY 79 VVDVSHEDPEVKFNWYDGVVEVHNNAKTKPREEQNSTYRVVSVTLVHLDHMLNKEYKC 138
Db 283 VVDVSEDDPDVQISWNNVEVLTATQTHREDYNSTLRVVSALPIQHQDWMGKEFKC 342
QY 139 KYSNKALPAPIEKTISKAKGQPREPQVYTLPPSRDELTKNOVSLTCLVKGYFSPDIAVEW 198
Db 343 KYNKALPAPIERTISKAGSRAPQVYTLPPSEEMTKKQVTLGCMVYDFMPEDIIYVW 402
QY 199 ESNQGPENNYKTPPPVLDSDGSFFLYSLKTVDSKRWQGNVFCSCVMHEALHNHYTKSL 258
Db 403 TNGKTELNYKTEPVLDSGDSGYFMYSLRVEKKNVENSYSYCSVWHEGLNHHHTKSF 462
QY 259 SLSPGK 264
Db 463 SRTPGK 468

RESULT 11
Q8R3H6 PRELIMINARY; PRT; 474 AA.
AC Q8R3H6;
DT 01-JUN-2002 (TRENBLrel. 21, Last sequence update)
DT 01-JUN-2002 (TRENBLrel. 21, Last sequence update)
DT 01-JUN-2002 (TRENBLrel. 21, Last annotation update)
DE Hypothetical 51.7 kDa protein.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RA Strausberg R.;
RL Submitted (MAR-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC025447; AAH25447.1;
DR MGI; MGI:2144967; A0044919.
DR InterPro; IPR003045; CytC_heme_bind.
DR InterPro; IPR003006; Ig_MHC.
DR Pfam; PF00047; Ig_3.
DR PROSITE; PS00190; CYTOCHROME_C; UNKNOWN_1.
DR PROSITE; PS00290; IG_MHC; UNKNOWN_1.
KW Hypothetical protein.
SQ SEQUENCE 474 AA; 51748 MW; 8608B57C6CD2874A CRC64;

Query Match
Best Local Similarity 56.4%; Score 806; DB 11; Length 474;
Matches 151; Conservative 44; Mismatches 59; Indels 177; Gaps 6;

QY 1 DWLKA----- 5
Db 54 NWKRRPGGLEWIGRIFFGDDTHYSGKFQKAKLTADKSSVTAFLQTLTSSDSAVY 113
QY 6 -----FYDKVAEKLKEAFD----- 20
Db 114 FCARDSYGDYD-----DNQCATVTVSSAKTTPPSVYPLAPCGDGTGSSVTL 163
QY 21 ---LKAFYDK-----VA 29
```

```
Db 164 GCLVKGYPEPVTWNSGSLSSVHTFPALQSLGYTMSSSVTVPSSTWPSQITVCSVA 223
QY 30 EKLKEAFMDK-----THTCPP-----CPAPELLGGPSVFLPPPKPKDTLMISRT 73
Db 224 HPASSTVDDKLEPSPGISTINPCPPCKBCHKCPAPNLEGSPSVIFPPNKKDVLMLSLT 283
QY 74 PEVTVVDVSHEDPEVKFNWYDGVVEVHNNAKTKPREEQNSTYRVVSVTLVHLDHMLNG 133
Db 284 PKVTCVVDVSEDDPDVQISWNNVEVLTATQTHREDYNSTLRVVSALPIQHQDWMGSG 343
QY 134 KEYCKVSNKALPAPIEKTISKAKGQPREPQVYTLPPSRDELTKNOVSLTCLVKGYFSPD 193
Db 344 KEFKCKVNNKLPSPGISTISKIKGLVRAPOVYILPPPAEQLSRKDVSLTCLVGVNPGD 403
QY 194 IAVVESNGQPPENNYKTPPPVLDSDGSFFLYSLKTVDSKRWQGNVFCSCVMHEALHNHY 253
Db 404 TSVENTSGHTKTEENTKDTAPVLDSDGSFYISKDKIKTSKWEKTSFSCNVRHEGLKNY 463
QY 254 TOKSLSLSPGK 264
Db 464 LKKTISRSPGK 474

RESULT 12
Q91Z05 PRELIMINARY; PRT; 473 AA.
AC Q91Z05;
DT 01-DEC-2001 (TRENBLrel. 19, Created)
DT 01-DEC-2001 (TRENBLrel. 19, Last sequence update)
DT 01-JUN-2002 (TRENBLrel. 21, Last annotation update)
DE Hypothetical 51.9 kDa protein.
GN AU044919.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RA Strausberg R.;
RL Submitted (JUL-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC010327; AAH10327.1;
DR MGI; MGI:2144967; A0044919.
DR InterPro; IPR003045; CytC_heme_bind.
DR InterPro; IPR003006; Ig_MHC.
DR Pfam; PF00047; Ig_3.
DR PROSITE; PS00190; CYTOCHROME_C; UNKNOWN_1.
DR PROSITE; PS00290; IG_MHC; UNKNOWN_1.
KW Hypothetical protein.
SQ SEQUENCE 473 AA; 51946 MW; CF625F008932AF12 CRC64;

Query Match
Best Local Similarity 56.0%; Score 801; DB 11; Length 473;
Matches 153; Conservative 44; Mismatches 66; Indels 156; Gaps 6;

QY 2 WLKAFYDK-----VAEKLKEAFD----- 19
Db 55 WVRQAPEKGLEWVAYNSGTTIYVADTVKGRFTISRDNAKNTFLQMTSLRSEDATAMY 114
QY 20 -----WL-----KAFYDK--- 27
Db 115 CARELWLRRIYWGQGTITVSSAKTTPPSVYPLAPGCGDGTGSSVTLGCLVKGYFPEV 174
QY 28 -----VAEKLKEAFMDK--- 39
Db 175 TVTWNSGSLSSVHTFPALQSLGYTMSSSVTVPSSTWPSQITVCSVAHPASSTVSKKL 234
QY 40 -----THTCPP-----CPAPELLGGPSVFLPPPKPKDTLMISRTPEVTVVDVSH 85
Db 235 EPSGPISITINPCPKCKBCHKCPAPNLEGSPSVIFPPNKKDVLMLSLTPKTVVVDVSE 294
QY 86 EDPEVKFNWYDGVVEVHNNAKTKPREEQNSTYRVVSVTLVHLDHMLNKEYCKKYSNKL 145
```



Db 2549 NVFVELSHSGIDVLWNFKDKETKPSKYYKIEAHGKIYKLTVLNMMKDDCKYFYAGE 2608  
Qy 10 ----- 9  
Db 2609 NITSGKLTVAGGAISKPLTDQTVAESQAEVCEPEVANPDSKGEWLRDGHLPNTNIRSE 2668  
Qy 10 ----- 9  
Db 2669 SDGHRRLIIAATKLDIDIGEYTKVATSKTSAKLKVEAVKIKTKLKNLTVTETQDAVTV 2728  
Qy 10 ----- 9  
Db 2729 ELTHPNVGVOMKNGVVLESNEKYAISVKGTIYSLRIKNCAYDES YGFRGLRGASA 2788  
Qy 10 ----- 9  
Db 2789 RLHVETVKIIKKPRDVTALENATVAFVSVDHVPVFWFKHVEIKPSDKHRLVSEKVV 2848  
Qy 10 ----- 9  
Db 2849 HKMLQNI SP DAGEYTA VVGOLECKAKLFVETLHITKTMK NIEVPETKTASFCEVSHF 2908  
Qy 10 ----- 9  
Db 2909 NVPSMWLKNVIEIEMSEKFI VOGK L HOLIIMNTSTEDSAEYTFVCGNDQVSATLVTP 2968  
Qy 10 ----- 9  
Db 2969 IMITSMKLDINAEKDTITFEVTYNYEGISYKWLKNGVEIKSTDKCQMRTKKLTHSLNIR 3028  
Qy 10 ----- 9  
Db 3029 NVHFGDAADYTFVAGKATSTATLYVEARHIEFRKHIDIKVLEKKRAMFECEVSEPDITV 3088  
Qy 10 ----- 9  
Db 3089 QMMDQELQITDRIKQEKYVHRLLIPSTRMSDAGKYTVVAGNVSTAKLFVEGRDVR 3148  
Qy 10 ----- 9  
Db 3149 IRSIKKEVQVIEKQAVVEFEVNEDDVDHMYKDGIEINFQVQERHKYVVERIHRMFIS 3208  
Qy 10 ----- 9  
Db 3209 ETRQSDAGEYTFVAGNRSSVTVLYVNAPEPPQVLQELQPTVQSGKPARFCAVISGRQP 3268  
Qy 10 ----- 9  
Db 3269 KISWYKEBQLLSTGFKCKFLHDGQEYTLILLIEAFPEDAAYTCEAKNDYGVATTASLSV 3328  
Qy 10 ----- 9  
Db 3329 EVPEVSPDQEMPVYPPIAITPLQDVTVSEGPQARFQCRVSGTDLKVSWSKDKKIKPSR 3388  
Qy 10 ----- 9  
Db 3389 FFRMTQPDYQLEIAEAYPEDEGTYTFVASNAGVQVSSSTANLSLEAPESILHERIQEI 3448  
Qy 10 ----- 9  
Db 3449 EMEMKEFSSFLSABEEGLHSAELQSKINETLELLSESPVYSTKFDSEKEGTGPIFIKE 3508  
Qy 10 ----- 9  
Db 3509 VSNADISMGDVATLSVTVIGIPKPIQVFPNGVLLTSPADYKFFVFDGDHSLIILFTKLE 3568  
Qy 10 ----- 9  
Db 3569 DEGEYTCMASNDYKTCISAYLKINSKGECHKDTETESAVAKSLEKLGCPPHFLKELK 3628  
Qy 10 ----- 9  
Db 3629 PIRCAQGLPAIFETVTVGEPAPTVPFKENKOLCTSVYTIHNPNGSGTFIVNDPQRED 3688

Qy 10 ----- 9  
Db 3689 SGLYICRAENMLGESTCAEALLVLEDDTMDTDPCKAKSTPEAPEDFPQTPLKGPVAEAL 3748  
Qy 10 ----- 9  
Db 3749 DSEQIATFVKDTILKAALITEENQOLSYEHIAKANELLSQLPGAQELQSILODKLTP 3808  
Qy 10 ----- 9  
Db 3809 ESTREFLCINGSIHQP LUKESP NLQLOI VOSQKTFSEKILMPEEPETQAVLSOTEKIF 3868  
Qy 10 ----- 12  
Db 3869 PSAMSIEQINSLSITVEPLKTL LAEP EGNYPQSSIEPPHMSYLTSAEEVLSPEKIVSDTN 3928  
Qy 13 ----- 12  
Db 3929 REQRTLQKQEAQSALILSQAEGHVESLQSPDVMISQVNYEPLVPSEHSCTEGGKILI 3988  
Qy 13 ----- 12  
Db 3989 ESAMPLENAGODSAVRIEIEGKSLRFP LAL EBEKQVLLKEEHSNDVMVPPDQIESKREPVA 4048  
Qy 13 ----- 12  
Db 4049 IKKVOEVQGRDLLSKESLLSGIP EQRNLN KIQICRALQA AVASEQPC L FSEWLRNIEKV 4108  
Qy 13 ----- 12  
Db 4109 EVEAVNITQEPRHIMCYLVTSAKSVTEVTIIIEDVDPQMANLKMELRDALCAIYEEI 4168  
Qy 13 ----- 12  
Db 4169 DILTAEGPRIQOGAKTSLQEEMDSFGSKVEPITEPEVESKYLSISPEVSYFNQSVRK 4228  
Qy 13 ----- 17  
Db 4229 YLDATPVYTKGVASAVVDEKQDESILKPSEKESESSESGTEEVATVKIQEAGGFIKEDG 4288  
Qy 18 ----- 17  
Db 4289 PMIHTPLVDTVSEGDIVHLTTSITNAKEVNNYFENKLVPSDEKFKCLQDQNTVTLIDK 4348  
Qy 18 ----- 17  
Db 4349 VNTEDHQGEYVCEALNDSGKTATS AKLTVVKRAAPVIRKRKIEPLEVALGHLAKFTCEIQS 4408  
Qy 18 ----- 23  
Db 4409 APNVRFQWFKAGREIYESDKCSIRSSKYISSLEILTRTQVDCGEYTCASNEYGSVSCA 4468  
Qy 24 ----- 23  
Db 4469 TLTVTEAVPPTFLSRPKSLTTFVGKAAKFICTVTGTPVIETIQWKDGAALSPSPNWKISD 4528  
Qy 24 ----- 23  
Db 4529 AENKHILELSNLTIQDRGVYSCASNKFGADICQAEILIIDKPHFIKIELEPVQSAINKKV 4588  
Qy 24 ----- 29  
Db 4589 HLEQVDEDRKVTVTWSKDGOKLPPGDKYKICFEKDIATLEIPLAKLKDSTYVCTASNE 4648  
Qy 30 ----- 29  
Db 4649 AGSSSCSATVTVREPPSFVKKVDPSYLMPLPGESARLHCKLKGPSVIOVTFWFKNNKELSE 4708  
Qy 30 ----- 29  
Db 4709 NTRMYFVNSEAILDITDKVEDSGSYSCAEAVNDVSDSCSTEIVIKPEPPSFIKLEPAD 4768

QY 30 ----- 29  
Db 4769 IVRGTNALLQCEYSGTGPPEISWFKDKQIRSSKKYRLFQKSLVCLEIFSPNSADGVEY 4828  
QY 30 ----- 29  
Db 4829 ECVVANEVKGCGMATHLLKEPPTFVKVDDLIAGLGGOTVTLQAAVRGSEPISVTWKMGQ 4888  
QY 30 ----- 29  
Db 4889 EVIREDGKIKMSFNGVAVLIIIPDVQISFGCKYTCLAEENAGSQTSVGELIVREPAKIIIE 4948  
QY 30 ----- 29  
Db 4949 RAEIIQVAGDPATLEVTVAGTPELPKWKYKDGRLPLVASKKYRISFKNNVAQLKFYSAEL 5008  
QY 30 ----- 29  
Db 5009 HDSGQYTFEISNEVGSSCETTFTVLDRDIAPFTFKPLRNVDVSVANGTCRLDCKIAGSLP 5068  
QY 30 ----- 29  
Db 5069 MRVSWFKDGEKIAASDRYIAFVEGTASLEIIRVDMNDAGNFTCRATNSVGKDSGALI 5128  
QY 30 ----- 29  
Db 5129 VQPPSFVTKPGSKDVLPGSAVCLKSTFGGSTPLTIRWFKGNKELVSGGSCYITKEALES 5188  
QY 30 ----- 29  
Db 5189 SLELYLVKTSDSGTCTCKVSNVAGVECSANLFVKEPATFVEKLEPSSLKKGDATQLAC 5248  
QY 30 ----- 29  
Db 5249 KVTGTPPIKITWFANDREIKESSKHRMSFVESTAVLRLTIDVGIEDSGEYCEAONEAGSD 5308  
QY 30 ----- 29  
Db 5309 HCSSIVIVKESPYTFKEFKPIEVLDVDMLLAEVAGTPPEITWFKDNTILRSGRKYKT 5368  
QY 30 ----- 29  
Db 5369 FIODHLVSLQILFVAADAGEYOCRVTVNEVGSSIC SARVTLREPPSFIIKIESTSSLRG 5428  
QY 30 ----- 29  
Db 5429 TAAFOATLKGSPLITVTLWLDSDTEITDDNIRMTFENNVAASLYLSGIEVKHDKGYVCOAK 5488  
QY 30 ----- 29  
Db 5489 NDAGIORCSALLSVKEPATITEEAVSIDVTQGD PATLOVRFSGTKEITAKWFKDGOELTL 5548  
QY 30 ----- 29  
Db 5549 GSKYKISVTDVTSILKIISTEKKDGEYTFEVQNDVGRSSCKARINVLDLIIPPSFTKKL 5608  
QY 30 ----- 39  
Db 5609 KKMDSIKGSFIDLECIAGSHPISIQWFKDDQEIASEKYKFSFHDNTAFLEISQLEGTD 5668  
QY 40 --THTC-----PP----- 45  
Db 5669 SGTYTCSATNKGHNQCSGHLTVKEPPYFVEKPKSQSDVNPNTRVQLKALVGGTAPMTIKW 5728  
QY 46 -----C----- 46  
Db 5729 FKONKELHSGAARSVKWDSTSTLELFAAKATDSGTICYCOLSNDVGTATSKATL FVEKPP 5788  
QY 47 ----PAP----- 49  
Db 5789 QFIKKPSVLVLRNGOSTTFECQITGTPKIRVSWYLDGNEITAIQKHGISFIDGLATFOI 5848  
QY 50 ----- 49

Db 5849 SGARVENSGTYVCEARNDACTASCSELKVKPEPTFIRELKPVVEVKYSDVECEVGTGT 5908  
QY 50 ----- 49  
Db 5909 PPFVETWLKNNREIRSSKKYTLTDRVSVFNLHITKCDPSDTGEYOCIVSNEGSCSCSTR 5968  
QY 50 ----- 49  
Db 5969 VALKEPPSPFIKKIENTTIVLKSSATFOSTVAGSPPI SITWLKDDQIILDEDDNVYISFVDS 6028  
QY 50 ----- 49  
Db 6029 VATLOIRSVDNHSGRYTQCAKNESGVERCYAFLLVQEPAQIVEKAKSVDTVTEKDPMTLE 6088  
QY 50 ----- 49  
Db 6089 CVVAGTPELVKWLKDGKQIVPSRYFSMPENNVPRIOSVMKQDSGXQYTFKVENDFGS 6148  
QY 50 ----- 49  
Db 6149 SSCDAYLRVLQDNI PPSTFKLT KMDKVLGSSIHMECKVSGSLPISAQWFKDGKEISTSA 6208  
QY 50 ----- 49  
Db 6209 KYRLVCHERSVSVLENNLEDTANYTCVSNVAGDDACSGILTVKEPPSFLVKPGRQOA 6268  
QY 50 ----- 59  
Db 6269 IPDSTVEFKAILKGTTPPFKIKWFKDDVELVSGPKCFITGLEGSTFSLNLYSVDASKTGOYT 6328  
QY 60 ----- 59  
Db 6329 CHVTNDVGSCTTMLLVTEPPKFVKLEASKIVKAGDSSRLCKIAGSPEIRVVWFRNE 6388  
QY 60 ----- 59  
Db 6389 HELPASDKYRMTFIDSVAVIQMNNLSTEDSGDFICEAQNPAGSTSCSTKVIVKEPPVFSS 6448  
QY 60 FPP----- 62  
Db 6449 FPIVETLKNAEVSELECELSGTPPFVVVYWKDKRQLRSSKKYKIASKNFHTSIHLNVD 6508  
QY 63 ----- 62  
Db 6509 SDIGEXHCKAQNEVGSDTCVCTVKLKEPPRFVSKLNSLTVVAGEPAELQASIEGAQPIFV 6568  
QY 63 ---KPKD-----TLM-- 69  
Db 6569 QWLKEKEVIRESENI RITFVENVATLQFAKAPANAGKYICQIKNDGGMEENATLMVL 6628  
QY 70 ----- 69  
Db 6629 EPAVIVEKAGPMTVTVGETCTLECKVAGTPELSVEWYKDGKLLTSSQKHKFSFYNKISSL 6688  
QY 70 ----- 69  
Db 6689 RILSVERQDAGTYTFQVQNNVGKSSCTAVVDVSDRAVPSPFTRLKNTGGVLGASCILEC 6748  
QY 70 ----- 69  
Db 6749 KVAGSSPISVAFHEKTKIVSGAKYQTFSDNVCTQLQNSLSDMDGNYTCVAANVAGSD 6808  
QY 70 ----- 69  
Db 6809 ECRAVLTVQEPSPFVKEPEPEPLVLPKNVTFTSVIRGTPPFKVNFRGARELVKGDRCNI 6868  
QY 70 ----- 80  
Db 6869 YFEDTVAELELNFNIDISQSGEYTCVSNAGQASCTTFLFVKPEAPALKRLSDHSVEPGK 6928  
QY 81 ----- 80

Db 6929 SIILESTYTGTLPISVTWKKGDNITTEKCNIVTTEKTCILEILNSTRKRDAGQYSCETE 6988  
QY 81 ----- 80  
Db 6989 NEAGRDVCGALYSTLEPPYFVTELEPLEANAGDSVSLQOVAGTPEITVSVWKGDTKLRP 7048  
QY 81 ----- 80  
Db 7049 TPEYRTYFTNNVATLVFNKNVINDSGEYTKAENSIGTASSKTVFRIQERQLPPSFARQL 7108  
QY 81 -----VDVSH 85  
Db 7109 KDIEQTGLVTLTCLRLNGSAPIQVCWYRDGVLLRDRDENLOTFSFDNVATLKILQTDLSH 7168  
QY 86 E----- 86  
Db 7169 SQYSCSASNPGLTASSSARLTAREPKKSPFFDIKPVSDIVJAGESADFECHVTGAQPMR 7228  
QY 87 ----- 86  
Db 7229 ITWSKDNKEIRPGGNYTITCVGNTPHLRILKVGKDSGYTCQATNDVGDKMCQAQLSVK 7288  
QY 87 -----DPEVK----- 91  
Db 7289 EPPKFVKKLEASKVAKQGESIQLECKISGSPEIKVSWFRNDSSELHESWKYNMFSINSVAL 7348  
QY 92 ----- 91  
Db 7349 LLINEASAEADSGDYICEAHNGVDGASCSTALTATVKAAPPVFTQKPSVPGALKGSDVILQCEI 7408  
QY 92 ----- 91  
Db 7409 SCTPPEVWVKDRKQVRNKKFKITSKHFDLSHLINLEASDVGEYHCKATNEVGSOTC 7468  
QY 92 ----- 91  
Db 7469 SCVAFKEPPFVVKKLSDTSTLIGDAVELRAIVEGFQPIVSVWLKDRGEVIRESENTRIS 7528  
QY 92 ----- 91  
Db 7529 FIDNIATLQGSPEASNSGKYICQKNDAGMRECSAVLTIVLEPARIIEKPEPMTVTGNP 7588  
QY 92 ----- 91  
Db 7589 FALECVVTGTPELSAKWFKDGRSADSKHITFINKVASLKI PCAEMSDKGLYSFEVKN 7648  
QY 92 ----- 91  
Db 7649 SVGKSNCVTSVHVSDRIVPPSFIRKLDVNAILGASVVLECRVSGSAPISVGWFDQGNEI 7708  
QY 92 ----- 91  
Db 7709 VSGPKQCSFSENVCNLNLSLEPSDGIYTCVAANVAGSDECSAVLTIVQEPPEQTPD 7768  
QY 92 ----- 91  
Db 7769 SVEVLPGHSLTFTSVIRGTPPKVWFKGSRRELYPGESCINISLEDFTVELELFEVQPLES 7828  
QY 92 ----- 91  
Db 7829 GDYSLVTNDAGSACSTTHLEFVKEPATFVKRLADFVETGSPIVLEATYTGTPPISVSWI 7888  
QY 92 ----- 91  
Db 7889 KDEYLISQSERCSITWTEKSTILESTIEDYAOYSCLIENRAGQDICEALVSVLEPPY 7948  
QY 92 ----- 91  
Db 7949 FIEPLEHVEAVIGEPATLQCKVDGTPEIRISWYKEHTKLSAPAYKMQFKNNVASLVINK 8008  
QY 92 ----- 91  
Db 8009 VHDSDVGEYSCKADNSVGAVASSAVLVITKERKLPFFPFARKLKDVKHETLGFVAFECRING 8068

QY 92 -----FNMVYDGV----- 99  
Db 8069 SEPLQVSWYKDGVLKKDDANLOTSEFVHNVTLIQILQTDOSHIGQYNCSASNPLGTASSSA 8128  
QY 100 ----- 99  
Db 8129 KLILSEHEVPPFDLKPVSVDLALGESGTFKCHVTGTAPIKITWAKDNREIRPGGNYKMT 8188  
QY 100 ----- 99  
Db 8189 LVENTATLTVLVKVGKDGAGQYTCYASNIAGKDCSAHLGVQEPFRFIKLEPSRIVKODE 8248  
QY 100 -----EVH 102  
Db 8249 FTRYECKIGGSPEIKVLWYKDETEIQESSKFRMSFVDSVAVLEMHNL SVEDSGDITCEAH 8308  
QY 103 NA----- 104  
Db 8309 NAAGSASSSTSLKVKEPPIFRKKPHPIETLKGADVLECELOGTPPFHVSWYKDKRELRS 8368  
QY 105 -----KTKPR----- 109  
Db 8369 GKYYKIMSENFILTSIHILNVDAADIGEYQCKATNDVGSOTCVGSIALKAPRFRVKKLSDI 8428  
QY 110 ----- 109  
Db 8429 STVVGKEVOLOTTIEGABPISVWPFKDKGEIVRESNIIWISYSENIIATLQFSRVEPANAG 8488  
QY 110 ----- 109  
Db 8489 KYTCQIKNDAGMOECFATLSVLEPATIVEKPESIKVTTGDTCTLECTVAGTPELSTKWPK 8548  
QY 110 ----- 109  
Db 8549 DGKELTSNDKYKISFPFNKVGSLKIINAPSDSGVYSFEVQNPVGKDSCTASLQVSDRTPV 8608  
QY 110 ----- 109  
Db 8609 PSFTRKLTNGLSGSSVVMCKYVGGSPPISVSWFHEGNEISSGKKYQTTLTDTNCTALTV 8668  
QY 110 -----EE----- 111  
Db 8669 NMLEESDGDYTCIATNMAGSDECSAPLTVREPPFVQKPDPMDLVTGTNTVFTTSIVKGT 8728  
QY 112 ----- 111  
Db 8729 PPFVSWFKGSSELVPGDRCNVLESDVALELFDVDTSQGEYTCIVSNEAGKASCTTH 8788  
QY 112 ----- 111  
Db 8789 LYIKAPAFVRLNDYSIEKGKPLILEGTFGTGTPPISVTWKKNGINVTGPSQRNITTTTEK 8848  
QY 112 -----QYN----- 114  
Db 8849 SAILLEPSSSTVEDAGQYNCYIENASGKDCSAQILILEPPYFVKQLEPVKVSVDGDSASIQ 8908  
QY 115 -----STYRV----- 119  
Db 8909 CQLAGTPEIGVSWYKGDTKLRPTTTYTKMHFRNNVATLVFNQVDINDSGEYICKAENSUGE 8968  
QY 120 -----VSV----- 122  
Db 8969 VSASTFLTVOEQKLPPFSRQLRDVQETVGLPVVFDCAISGSEPISVSWYKDGKPLKDSP 9028  
QY 123 ----- 122  
Db 9029 NVQTSFLDNTATLNIFKTDRLAGQY SCTATNPIGSASSSARLILTEGKNPPFFDIRLAP 9088  
QY 123 -----LTVLHQD----- 129  
Db 9089 VDAVVGESADFECHVTGTQPIKVSMAKDSRISRGKGYQISYLENSAHLTVLKVDDKDGSG 9148

```
QY 130 -----WLNKGE----- 135
Db 9149 QYTCYAVNEVGKDSCTAQLNIKERLIPPSFTKRLSETVETEGNSFKLEGRVAGSQITV 9208
QY 136 -----YKCVSN----- 142
Db 9209 AWYKNNIEIOPTNCEITFKNNTLVQLVRKAGMNDAGLYTCKVSNDSAGSALCTSSIVIKE 9268
QY 143 ----- 142
Db 9269 PKKPPVFDQHLTPVTYSEGEVQVLSCHVQSEPIRQWLKAGREIKPSDRCSFSFASGTA 9328
QY 143 ----- 142
Db 9329 VLELRDVAKADSGDYVCASNAVAGSDTTKSKVTIKDKPAVAPATKKAANDGRLLFFVSEPQ 9388
QY 143 ----- 142
Db 9389 SIRVEKTTATFIKVGGDPIPNVKWTGKWROLNOGGRVFIHQKGDEAKLEIRDTTKTD 9448
QY 143 ----- 142
Db 9449 SGLYRCVAFNEHGEIESNVNLQVDERKKQEKIEGDLRLMLKKTPIILKKGAGEEEDIME 9508
QY 143 ----- 142
Db 9509 LLKNVDPKEYEKYARMYGITDRGLLQAFELLKQSOBEETHRLEIEIERSERDEKEFEE 9568
QY 143 ----- 142
Db 9569 LVSEIQORLSQTEPVTVLKDIENQTVLKNDVAFEDIKINYPEIKLSWYKGTLEKLEPSD 9628
QY 143 ----- 142
Db 9629 KFEISIDGRHLRVKNCQLKDOGNVLCVCGPHIASAKLTVIEPAWERHLQDVTLKEGQT 9688
QY 143 ----- 142
Db 9689 CTMTQCFVSNVSKSEWFRNGRILKPOGRHKTEVEHKVHLTIADVRAEDQGYTCRYEDL 9748
QY 143 ----- 142
Db 9749 ETSAEILRIEAPIQFKRIONIVVSEHOSATFECEVSFDDAIVTWYKGPTELTESOKYNF 9808
QY 143 ----- 142
Db 9809 RNDGRCHYMTIHNWTPDEGVSVIARLEPRGEARSTAEILYLTKEIKLEKLPDPIDSR 9868
QY 143 ----- 142
Db 9869 VPIPTMPIRAVPPPEIPVAVPPIPLLLPTPEEKPPPKRIEVTKKAVKDKAKKVAKPK 9928
QY 143 ----- 142
Db 9929 EMTPREIEVKPPPTTLIPAKAPEIDVSSKAEEVKIMTITRKKVEQKEAVEKKQA 9988
QY 143 ----- 142
Db 9989 VHKEKRVFIESPEPYDELEVEPYTEPFEOYYVEEDPEDEYEEKVKAKEVHEEWEEDFE 10048
QY 143 ----- 142
Db 10049 EGOEYTEREGYDEGEWEAEAYOEREVIQVQKEVYESHERKVPKVPKAPKPPPKVI 10108
QY 143 ----- 142
Db 10109 KKPVIKIEKTSRMBEEKVQVTKVPEVSKKIVPQKPSRTPVOEEVIEVKVPAVHTKKMV 10168
QY 143 ----- 142
Db 10169 ISEKMFASHTEEEYSVTVPQKEIVTEEKIHVAVSKRVEPPPKVPELPEKPAPEVA 10228
QY 143 ----- 142
```

```
Db 10229 PVPIPKKVEPPAPKVPVEVPKKVPVEKKVPVPKKEPAAPPKVPVEPKKVPVEEKIPVPV 10288
QY 143 ----- 142
Db 10289 AKKKEAPPKVPVEVQKRVVTEEKITIVTQREESPPPAVPEIPKKKVPVEERKVPVKKEEV 10348
QY 143 ----- 142
Db 10349 PPPKVPALPKVPVEEKVAVVPVPAKAPPPRAEVSKTIVVEEKRFVAEKLFSFVAPQR 10408
QY 143 ----- 142
Db 10409 VEVTRHEVSAEEEWSYSEEEEGVSISVYREEREEREEAEVTEYVMEPEEYVVEEKLH 10468
QY 143 -----KALPAPI----- 149
Db 10469 IISKRVEAEPAEYTEROEKKIVLKPPIAKIEEPPPAKVPPEAKKIVPEKKVPAPVPKKE 10528
QY 150 ----- 149
Db 10529 KVPPPKVPPEEPKVPVEKVPVKVPMKEEPLPAKVTERHMQITQEEKVLVAVTKKEAPPK 10588
QY 150 ----- 149
Db 10589 ARVPEEPKRAVPEEKVLKPKREEPAPKVTETFRKRWKEEKVSIAPKREPOPIKEVT 10648
QY 150 ----- 149
Db 10649 IMBEKERAYTLEEAHSVQREEEYEEYDYKEFEYETPEYDQYEEYEEYEEYEE 10708
QY 150 ----- 149
Db 10709 HEEYITEPEKPIPVKVPVEEPVTKPAKPAKVLKKAPEEKVPVPIPKLKLPPPKVPE 10768
QY 150 ----- 149
Db 10769 EPKKVFEKIRISITRKEQVTEPAKVPKPKRVAAEEKVPVPRKEVAPPVPRVPEVPK 10828
QY 150 ----- 149
Db 10829 ELEPEEVAFESEVVTHVEEYLVVEEEYIHEEEFITEEEVVVPIPVKVPVEPKKVPPEE 10888
QY 150 ----- 149
Db 10889 KKPVPVKKKEAPPKVPVEVPKKPEEKVPVLIIPKKEKPPPAKVPVEPKKVPVEEKVPVPV 10948
QY 150 ----- 149
Db 10949 PKKVEAPPKVPVEVPKVPVEKVPVPVPAKPKKVEAPPKVPVEPKKLIPEEKKPTVPKPV 11008
QY 150 ----- 149
Db 11009 EAPPKVPKAREVPVPVVALPQEEVLFEEIIVPEEVLPEEEVLPEEVLPEEVLPEEVL 11068
QY 150 ----- 149
Db 11069 PEEETIPPEEEVPPPEEYVPEEEFVPEEVLPEVKVPVPVPAVPEIKKKVTEKKVVI 11128
QY 150 ----- 149
Db 11129 PKKEAPPKVPVEVPKVKKEKRIILPKEEVLVPVEVTEPEEPISEETIPEEPPSIEEV 11188
QY 150 ----- 149
Db 11189 EEVAPPRVPEVIKKAPEAPTVPKVPKVEAPPKVKSKKIPKEKVPVPVOKKEAPPKVPVEV 11248
QY 150 ----- 149
Db 11249 PKKVPKKVLVPKKEAVPPAKGRTVLEEKYSVAFRQEVVVVKERLELEWVAEVEEPIEEE 11308
QY 150 ----- 149
```

Db 11309 EFHEVEYFEGEFHEVEBFIKLEOHRVEEHRVEKHVRVIEVFEAEVEVEFEKPKAPPK 11368  
QY 150 ----- 149  
Db 11369 GBEISEKIIIPPKPPKPPKVVPRKEPPAKVPEVKKIIVVEKVRVPEPRVPPTKVDPDLPP 11428  
QY 150 ----- 149  
Db 11429 KEVVPKKVPVPAKKEAPPKVPAPKEVVPPEKKVPVPPPKKPEVPTKVPEVPAKAAV 11488  
QY 150 ----- 149  
Db 11489 PEKKVPEAIIPKPEPPPEAPKEVPEKKVPAAPPKKPEVTPVKVPEAPKEVVPPEK 11548  
QY 150 ----- 149  
Db 11549 VPPPPKKPEVPTKVPEVKVAVPEKKVPEAIIPKPEPPPEVPEPEPEVALEPPAEV 11608  
QY 150 ----- 149  
Db 11609 VEEPEAAPQVTPPKKPVPEKKAPAVAKKPELPPVKVPEVPEKKVPLVVPKK 11668  
QY 150 ----- 149  
Db 11669 PEAPPAKVPPEVKVPEKKVAVPKKPEVPPAKVPEVKKPVLEKPAVVPPEAESEPPP 11728  
QY 150 ----- 149  
Db 11729 EYVEPEETAPEETAPEBEKVPVVAEBEPEVPPPAVPEEPKIIPEKKVPVIRKPEAP 11788  
QY 150 ----- EKTISK ----- 155  
                  |||  
Db 11789 PKPEPEKVIKPKLPPPPPPPPAPPKEDYKIFQLKAIPKKKVPPEKQVPEKVELT 11848  
QY 156 ----- 155  
Db 11849 PLKVPGEKKVRKLLPERKPEPEEYVLKSVLRKPEEPEKPEKLEKVKKPAVPEPP 11908  
QY 156 ----- 155  
Db 11909 PPKPVEVEVPTVKRERKIPEPTKVPEIKPAIPAPPEPKKPEAEVKTIKPPVPEPEP 11968  
QY 156 ----- 155  
Db 11969 TPIAAPVTVPVVGKAAKAPKEAAKPKPIKVPKKTSPTEAERRKLKRGSGKEPP 12028  
QY 156 ----- 155  
Db 12029 DEAPFTYQLKAVPLKFVKEIKDIILTESEFVGSSAIFECVSPSTAITTMMKDGSNIRE 12088  
QY 156 ----- 155  
Db 12089 PKHRTADGKDKRLHIDVQLSDAGEYTCVLRGNKEKTSTAKLVVEELPVPRFKTLEEE 12148  
QY 156 ----- AKGQP ----- 160  
                  ||||  
Db 12149 VVVWKGPLYLSCELNKERDVVWRKDKIVVEKPGRIVPGVIGLMRALTINDADDTDACT 12208  
QY 161 ----- 160  
Db 12209 YTVTVENANNLECCSVKVVVEIRDWLKPIRDQHVKPKGTAFACDIAKDTPNIKWFKG 12268  
QY 161 ----- 160  
Db 12269 YDEIPAEPNDKTEILRDGNHLYLKIKNAMPEDIAEYAVEIEGKRYPAKLTIGEREVELLK 12328  
QY 161 ----- 160  
Db 12329 PIEDVTIYESASFDAEISEADIPQWKLKCELLRPSPTCEIKAEGGKRLTLRAKVLID 12388  
QY 161 ----- 160  
Db 12389 QAGEVLYQALNAITTAITLTVKEIELDFAVPLKDVTPERRQARFECVLTRANVTWSKGP 12448

QY 161 ----- 160  
Db 12449 DIIKSSDKFDIIADGKKHILVINDSQFDDGEGVYTAEBEGKTSARLFTVIGIRLKFNMSPLE 12508  
QY 161 ----- 160  
Db 12509 DQTVKEGETATFVCESLSEKMHVWFKNDAKLHTSRTVLJSSGKTHKLEMEVTLDDLS 12568  
QY 161 ----- 160  
Db 12569 QIKAQVKELSTAQLKVLADPYFTVKLHDKTAVEKDEITLKCEVSKDVPVWFKDGBEI 12628  
QY 161 ----- 160  
Db 12629 VPSPKYSIKADGLRILKIKKADLKDKGVEYVCDGTDKTKANVTVEARLIKVEKPLVGYE 12688  
QY 161 ----- 160  
Db 12689 VVGETAHEFIEIELSEBPDVHGQWKLKGQPLTASPDCEIIEDGKKHILILHNCQLGMTGEYS 12748  
QY 161 ----- REPQ ----- 164  
                  |||  
Db 12749 FOAANAASANKVKELPLIFITPLSDVKVFEKDEAKFECEVSREPFTRWLKGTOEITG 12808  
QY 165 ----- 164  
Db 12809 DDRFELIKDGTKHSVVIKSAAFEDAKYFEAEDKHTSGKLLIEGIRLKFPLKDVATAK 12868  
QY 165 ----- 164  
Db 12869 EKESAVFTVELSHDNIRVKWFKNDQRLHTRSRSVMDQEGKTHSITFKDLSIODTSIRVE 12928  
QY 165 ----- 164  
Db 12929 AMGMSSEAKLTVLEGDVPYFTGKLQDYTGVEKDEVILQCEISKADAPVWFKDGEIKPSK 12988  
QY 165 ----- 164  
Db 12989 NAVIKADGKKRMLILKALKSDIGQYTCDCGTDKSGKLDIEDREIKLVRPLHSVEVNET 13048  
QY 165 ----- 164  
Db 13049 ETARFETEISDDIHANWKLKGEALLQTPDCEIKBEGKIHSLVHLHNCRLDQGTGVDFQAA 13108  
QY 165 ----- 164  
Db 13109 NVKSSAHLRVKPRVTIGLLRPLKDVTVTAGETATFDCELSYEDIPIVEWYLGKKLEPSDKV 13168  
QY 165 ----- 164  
Db 13169 VPRSEKVTTLTRDVKLEDAGEVOLTAQDKFTHANLFFVKPEPPVFTKPLEDQTVEEGAT 13228  
QY 165 ----- 164  
Db 13229 AVLECEVSRENKAKVWFKNGTEILKSKKYEIVADGRVRKLVIHDTCPEDIKTYTCDAKDF 13288  
QY 165 ----- 164  
Db 13289 KTSCLNINVPPHVEFLRPLTLQVREKEMARPECELSRENKAKVWFKDGAIEKKKKYDI 13348  
QY 165 ----- 164  
Db 13349 ISKGAVRILVINKLLDDEAEYSCEVRTARTSGMLTVLEEEAVFTKNLANIEVSETDTIK 13408  
QY 165 ----- 164  
Db 13409 LVCEVSKPGAIEVIWYKGDDEEIIETGRYEILTEGRKRILVIONAHLEDAGNVCRLPSSRT 13468  
QY 165 ----- 164  
Db 13469 DGVKVVHELAAEFISKPONLEILEGEKAEFVCSISKSEFPVQWKRDDKTLESGDKYDVIA 13528



QY 165 ----- 164  
Db 13529 DGKRVLVVKDAFLQDMGYVVMVGARAAAHLTVEKLRIVVPLKDRVKEQEVFNC 13588  
QY 165 -----VYTL----- 168  
Db 13589 EVNTEGAKAKFRNEAIFDSSKYIILQKDLVYTLRIRDAHLDDQANYVSLNHRGENV 13648  
QY 169 ----- 168  
Db 13649 KSAANLIVEEDLRIVEPLKDIETMEKKSVTWCNKVRLNVLKWTNGEEVFFDNRVSY 13708  
QY 169 ----- 168  
Db 13709 RVDKYRKMILTICDGPDEGEYIVTAGQKSVAAELLIIEAPTEFVEHLEDQVTFEDDAV 13768  
QY 169 ----- 168  
Db 13769 FSCQLSREKANVKWYRNGREIKGKKYKFKDGSIHRLIIKDCLDDECEYACGVEDRS 13828  
QY 169 ----- 168  
Db 13829 RARLFVEEIPVEIIRPPQDILEAPGADVFLAELNKKDVEQVLRNMMVVVQDCKHOMS 13888  
QY 169 ----- 168  
Db 13889 ECKIHRLOICDIKPRQGEYRFIAKDEARAKLELAAAPKIKTADQDLVVDVGKPLTMV 13948  
QY 169 ----- 168  
Db 13949 PYDAYKAEAEWFKENEPLSTKTIDTTAEQTSFRILEAKKGDKGRYKIVLQNKHKAEGF 14008  
QY 169 ----- 168  
Db 14009 INLKVIDVPGVVRNLEVTFETFDGEVSLAWEELTDGSKLIIGYVVERRDIKRTWVLAD 14068  
QY 169 ----- 168  
Db 14069 RAESCEFTVTGLQKGVEYLFVRSARNRVGTGEPVETDNPVEARSKYDVPGLNVTID 14128  
QY 169 ----- 168  
Db 14129 VNRFGVSLTWEPPEYDGGAEITNYVIELRDKTSIRWDATMTVRAEDLSATVTDVVEGEY 14188  
QY 169 ----- 168  
Db 14189 SFRVRAQNRIGVGKPSAATPFVKVADPIERPSPVNLTSDDQTSVQLKWEPPPLKDGGS 14248  
QY 169 ----- 168  
Db 14249 PILGYIIERCEEKGNWIRCNMMLVPELITYKVTGLEKNKLYRVSANENKAGVSDPSEIL 14308  
QY 169 ----- 168  
Db 14309 GPLTADDAFVEPTMDLSAFKDGLEIVPNPITILVPSTGYPRPTATWCFGKVLGTDRV 14368  
QY 169 ----- 168  
Db 14369 KMTLSAYAEVLISPSERSDKGIYTLKLENRVKTIISGEIDVNVIAIPSAPKELKFGDITK 14428  
QY 169 ----- 168  
Db 14429 DSVHLTWEPDDGGSLTGCVVVEKREVSRTWTVMDFVTDLEFTVPLVQGEYLFKV 14488  
QY 169 ----- 168  
Db 14489 CARNKCGPGEPAYVDEPVNMSTPATVPDPPENVKWRDTANSIFLTWDPKNDGGSRIK 14548  
QY 169 ----- 168  
Db 14549 YIVERCPRGSKWVACGEPVAETKMEVTGLEBGKWYAYRVKALNRQAGASKPSRPTIEIOA 14608  
QY 169 ----- 168

Db 14609 VDTQEAEIFLDVYKLLAGLTVKAGTKIELPATVTKPEPKITWTKADMILKODKRITIE 14668  
QY 169 ----- 168  
Db 14669 VPKKSTVTIVDSKRSDTGTYYIEAVNVCGRATAVVEVNVLDKPGPPAAFDITDVTNESCL 14728  
QY 169 ----- 168  
Db 14729 LTNWPRDDGGSKITNVVVERRATDSEVWHKLSSTVKDTNFKATKLIIPNKEYIFRVAAN 14788  
QY 169 ----- 168  
Db 14789 MYGVGEVQASPIAKYQFDPGPTRLPESDITKDAVTLTWCEPDDGSGSPITGYWVER 14848  
QY 169 ----- 168  
Db 14849 LDPTDKWVRCKMPPKVDTTYRVKGLTNKKYRFRVLAENLAGPGKPSKSTEDILIKDPI 14908  
QY 169 ----- 168  
Db 14909 DPPWPFKPTVKDVGKTSVRLNWTKPEHDGGAIESVVIEMLKTGTDWVRVAEGVPTTQ 14968  
QY 169 ----- 168  
Db 14969 HLLPGLMEGOEYSFRVRAVNAKAGESEPSDPVLCREKLYPPSPRWLEVINITKTAD 15028  
QY 169 ----- 168  
Db 15029 LKWTVEKGGGSPITNVIYEKRDVRRKGQVTDVTKDKTCTVPLTEGSLYVFRVAAN 15088  
QY 169 ----- 168  
Db 15089 AIGQSDYTEIEDSVLAKDTFTTTPGPPYALAVVDVKRHVDLKWEPPKNDGGRPIQRYVIE 15148  
QY 169 ----- 168  
Db 15149 KKERLGRWYKAGTAGPCNFRVTDVIEGTEVQFQVRAENEAGVGHSPSEPTILSIEDP 15208  
QY 169 ----- 168  
Db 15209 TSPSPPLDLHVTDAGRKHIAIAWKPEKNGGSPITGYHVEMCPVGTCKMVRVNSRPIKD 15268  
QY 169 ----- 168  
Db 15269 LKFKVEEGVVPDKKEYVLRVRAVNAIGVSEPSSEISENVVAKDPCKPTIDLETHDIIIVIEG 15328  
QY 169 ----- 168  
Db 15329 EKLSIPVPFRAPVPVTVSMHKGKVKASDRLTMKNDHISAHLEVPKSVRADAGIYITL 15388  
QY 169 ----- 168  
Db 15389 ENKLSGATASINVKVIGLPGCPCKDIKASDITKSSCKLTWEPPEFDGTPILHYVLERREA 15448  
QY 169 ----- 168  
Db 15449 GRRTYIPVMSGENKLSWTVKDLIPNGEYFFRVRKAVNKVGGEYIELKNPVIADPKQPPD 15508  
QY 169 ----- 168  
Db 15509 PPVDVEVHNPTAEAMTITWKPPLYDGGSKIMGYIIEXIAKGEERKRCNEHLVPLTYTA 15568  
QY 169 ----- 168  
Db 15569 KGLECKEYQFVRAENAGISEPSRATPPTKAVDPIDAPKVILRTSLEVKRGDEIALDA 15628  
QY 169 ----- 168  
Db 15629 SISGSPYTTITWIKDENVIYVPEETKKRAAPLVRRRKGEVQEEFPVPLTQLRLSIDNSKK 15688  
QY 169 ----- 168

Db 15689 GBSQLVRDSLRPDHGLYMIKIVENDHGIAPCTVSLDTPGPNFVFDIRKTSVLCK 15748  
QY 169 ----- 168  
Db 15749 WEPPLDDGGSEIINTLEKKDKTKPDSEWIVVTSLRHCKYSVTKLIEGKEYLFRVAEN 15808  
QY 169 ----- 168  
Db 15809 RFGPGPCVSKPLVAKDPGPPDAPDKPIVEDVTSNSMLVKWNEPKDNGSPILGYWLEKR 15868  
QY 169 ----- 168  
Db 15869 EYNSTHWSRVNKSLLNALKANVDGLLEGLTYVFRVCAENAAGPKFSPDPKTAHDPTS 15928  
QY 169 ----- 168  
Db 15929 PPGPPIPRVDTSSSTTIELEWEPFAPNGGGEIVGYFVDKLVGTNWSRCTEKMVKRQY 15988  
QY 169 ----- 168  
Db 15989 TVKEIREGADYKLRSVAVNAAGEPPGETQPVTVAEPOEPPAVELDVSVKGGIIMAGKT 16048  
QY 169 ----- 168  
Db 16049 LRIPAVVTGRPVPTKVTWKEGELDKDRVVDNVGTSKSELIIKDALKRDHGRVITATNS 16108  
QY 169 ----- 168  
Db 16109 CGSKFAARVEFDPGPPVLDLKPVTNRKMCLLNWSDEDDGGSEITGFIIERDKAMH 16168  
QY 169 ----- 168  
Db 16169 TWRQPIETERSKCDITLLEGOEYKFRVIANKKFGCPPEIGPILAVDPLGPPTSPEL 16228  
QY 169 ----- 168  
Db 16229 TYTERKSTITLDWKEPRNSGSPIOGYIIKKRRHDKPDFERVNKRKLCPTTSFLVENLDE 16288  
QY 169 ----- 168  
Db 16289 HQMEYFRKAVNEIGESPSPLNVVIOQDEVPPTIKLRSVRGDTIKVKAPEVHIAPD 16348  
QY 169 ----- 168  
Db 16349 VTGLPMPKIEWSNETVIEKPTDALQITKEEVSSEAKTELSIPKAVREDKGYTVVTASN 16408  
QY 169 ----- 168  
Db 16409 RLGSVFRNVHVEYDRPSPPRMLAVTDIKAESCYLTWDAPLDNGGSEITHYVIDRKDASR 16468  
QY 169 ----- 168  
Db 16469 KKAEEVNTAVKRYGIWKLIPNGQYFRVAVNKKYGISDECKSDKVVIOQPYELPGP 16528  
QY 169 ----- 168  
Db 16529 PGKPKVLARTKGSMLVSWTPPLDNGSGPITGYWLEKREBEGSPYWSRVSRAPIITKVGLKGV 16588  
QY 169 ----- 168  
Db 16589 EENVPRLLGVKYQPRAMAINAAGIGPPSEPSDPEVAGDPIFPGPPSPCEVKDKTSSI 16648  
QY 169 ----- 173  
Db 16649 SLGNWPPAKDGGSPIKGYIVEMQEBGTTDWKRVNEPDKLITTCCEVVPNLKELRYFRV 16708  
QY 174 ----- 173  
Db 16709 KAVNEAGESEPSDITGTGIEPATDIOBEPEVFDIGAQDCLVCKAGSOIRIPAVIKGRPTPK 16768  
QY 174 ----- 173  
Db 16769 SSWEPDGAKKAMKGVHDIPEDAQLETAENSSVIIPECKRSHTGKYSITAKNKAGOKT 16828

QY 174 ----- 173  
Db 16829 ANCRVKVMDVPGPPKDLKVSDITRGSCLSWKMPDDDDGDRIKGYVIEKRTIDGKAWTKV 16888  
QY 174 ----- 173  
Db 16889 NPDCGSTTFVVPDLLSEQQYFFRVRRAENRFGIGPPVETIQTARTARDPIYPPDPPIKLGIG 16948  
QY 174 ----- 173  
Db 16949 LITKNTVHLSWKPPKNDGGSPVTHIVECLAMDPTGTTKEANRQCNKRQDVEELQFTVEDL 17008  
QY 174 ----- 173  
Db 17009 VEGGEYFRVKAUNAAGVSKPSATVGPCDCQRDPMPPSIDLKEFEVEBEGTWNVINVAKIK 17068  
QY 174 ----- 173  
Db 17069 GVPFPTLTWFKAPKKPDNKEPVLVDYTHVNLVVDCTFLVIPQSRSDTGLYTITAVNN 17128  
QY 174 ----- 173  
Db 17129 LGTASKEMLNVLGRPPPVGPIKFSVSADQMTLSWPPPKDDGSGKITNYVIEKREANR 17188  
QY 174 ----- 173  
Db 17189 KTWVHVSSEPRECTYTIKLLLEGHEVYPRIMAQNKYIGIGEPDLSBEPETARNLFSVPGAPD 17248  
QY 174 ----- 173  
Db 17249 KPTVSSVTRNSMTVNWBEPEYDGGSPVTGYWLEMDKTTSKRWKVRNRPDIKAMTLGVSYK 17308  
QY 174 ----- 173  
Db 17309 VTGLIEGSDYQFRVYAINAAGVGPASLPSPATARDPIAPGPPPKVTDMTKSSADLEW 17368  
QY 174 ----- 173  
Db 17369 SPPLKGGSKVTGYIVEYKEEKEWEKDKVEKGTCLVVTGLKEGAFYKFRVSAVNITA 17428  
QY 174 ----- 173  
Db 17429 GIGEPGEVTDVTEMKDLVSPDLQLDASVRDRIVVHAGGVIIRIIAYVSGKPPPTVTWNMN 17488  
QY 174 ----- 173  
Db 17489 ERTLPOEATIIETAISSSWIKNCORSHQGVYSLAKNEAGERKKTIIIVDLDVPGPVCT 17548  
QY 174 ----- 173  
Db 17549 PFLAHLNLTNESCCLTWFSPEDDGGSPITNYVIEKRESDRRAWTPVTYIVTRONATVOGLJ 17608  
QY 174 ----- 183  
Db 17609 QKAYFFRIAAENSIGMGPFVETSEALVIREPITVPERPEDLEVKEVTKNTVTLTWNPBK 17668  
QY 184 ----- 183  
Db 17669 YDGGSEIINYLESLIGCTEKFKVTNDNLNLSRKYTVKLGEGDTYFRVSAVINVGOK 17728  
QY 184 ----- 183  
Db 17729 PSFCTKPTCKDELAPPTLHLDLFRDKLIRVCEAFALTGRYSGKPKPKVSWFKDEADVLE 17788  
QY 184 ----- 189  
Db 17789 DDRTHIKTTPATLALEKIKAKRSDSGKYCVVVENSTGSRKGFCQVNVVDRPVPVGVVSF 17848  
QY 190 ----- 189  
Db 17849 DEVTKDYMVISWKPPLDDGGSGKITNYIIEKKEVKDVMVMTSASAKTCKVSKLLEGKD 17908

QY 190 ----- 189  
Db 17909 YIFRIHAENLYGISDPLVSDSMKAKDRFVPDAPQPIVTEVTKDSALVTWNKPHDGGKP 17968  
QY 190 ----- 189  
Db 17969 ITNYILEKRETMKSRWARVTKDPIHPYTKFRVPDLLEGQYEFVSAENEIGIDFSPPS 18028  
QY 190 ----- 189  
Db 18029 KPVFAKDPIAKPPVNPPEALDTCNSVDLTWOPPRHDGSKILGIYIVEYKVGDEWRR 18088  
QY 190 ----- 189  
Db 18089 ANHTPESCPEYKYVTGLRDGQTYKFRVLAVNAAGESDPAHPVPEVLVKDRLEPPELILD 18148  
QY 190 ----- 189  
Db 18149 ANMAREOHIKVGDTLRLSAIIKGVPPPKVTWKEDRDAPTKARIDVTPVGSKEIRNAAH 18208  
QY 190 ----- 189  
Db 18209 EDGGIYSLTVENPAGSKTVSKVLVLDKPGPRDLEVSEIRKDCSYLTWKEPLDDGSGVI 18268  
QY 190 ----- 189  
Db 18269 TNYVVERRDVASQWSPLSATSKKKSHFAKHLNEGNOYLFRAAENQYGRGPFVETPKPI 18328  
QY 190 ----- 189  
Db 18329 KALDPLHPPGPKDLHHVDKTEVSLVWNKPRDRDGGSPITGYLVEYQEGTQDWIKFKT 18388  
QY 190 ----- 189  
Db 18389 VTNLECVVTGLQOGKTYFRVRAENIVGLGLPDTTPIEQOEKLVPPSVELDKLIEGLV 18448  
QY 190 ----- 189  
Db 18449 VKAGTTVRFPAILIRGVVPPTAKWTTDGGSEIKTDEHYTVETDNFSSVLTIKNCLRRDTGEY 18508  
QY 190 ----- 189  
Db 18509 QITVSNAAKSTVAVHLTVLDVPGPTGPIINILDVTFEHTISWQPPKDDGGSPVINYIV 18568  
QY 190 ----- 189  
Db 18569 EKQDTRKDTWGVSSGSKTKLIPHLOKGEYVFRVRAENKICGVGPPLDSTPTVAKHKF 18628  
QY 190 ----- 189  
Db 18629 SPPSPGKPVVTDITENAAATVSWTLPKSDGGSPITGYIMERREVTKGWRVKNKPTIADLK 18688  
QY 190 ----- 189  
Db 18689 FVRTGLYEGNTYEFVRAENLAGLSPSSDPIKACRPKPPGPPINPKLKDOKSRETAD 18748  
QY 190 ----- 189  
Db 18749 LVWTKPLSDGSPILGVVVEQCPGTAQWNRINKDELIROCAFVPCGLIEGNEYRFRKA 18808  
QY 190 ----- 189  
Db 18809 ANIVGEPELAEVIAKDILHPPEVELDVTCDVITVRVGTIRILARVKGRPEPDIT 18868  
QY 190 ----- 189  
Db 18869 WTKGKVLVREKRVLDLIQDLPRVELQKEAVRADHKGYYIISAKNSSGHAOGSAIVNVLD 18928  
QY 190 ----- 189  
Db 18929 PGPCQNLKVTNVTKENCTISWENPLDNGSGSEITNFIVEYRKPQKQGSIVASDVTKKLK 18988  
QY 190 ----- 189

Db 18989 ANLLANNEYFRVCAENKVGPGTIETKTPIALAINPIDRGPENLHIADKGKTFVYLKW 19048  
QY 190 ----- 189  
Db 19049 RRPDYDGGSPNLVSHVFERRLKGSDDWERVHKSGIKETHYMWDRCVENQIYEFVQTKNEG 19108  
QY 190 ----- 189  
Db 19109 GESDWVKEEYVVKEDLQKPVLDLKLSGVLTVKAGDTIRLEAGVGRGKPFPEVAMTKDKDA 19168  
QY 190 ----- 189  
Db 19169 TDLTRSPRVKIDTRADSKFSLTAKAKSDGKYVVTATINTAGSFVAYATVNVLDKPGFVR 19228  
QY 190 ----- 189  
Db 19229 NLKIVDVSSDRCTVCWDDPPEDDGGCEIQNYILEKCEIKRMWSTYSATVLTGPGTTVTRLI 19288  
QY 190 ----- 189  
Db 19289 EAGNEYIFRVRAENKIGTGPPTESKPVIATKYDKPGRDPPEVTKVSKEEMTVVWNPPEY 19348  
QY 190 ----- 189  
Db 19349 DGGKSITGYEFLEKKEKHSRWPVNKSAIPERRMKVONLLPDHEYQFRVKAENEIGICEP 19408  
QY 190 ----- 189  
Db 19409 SLSPRPVAKDPIEBPPGPPNFRVVDTKHSITLWGKGPVYDGGAPIIGYVVMRPKIAD 19468  
QY 190 ----- 189  
Db 19469 ASPDEGKRCNAAQLYRKEFTVTSLDENQEYEFVCAQNOVGIGRPAELKEAIKPREIL 19528  
QY 190 ----- 189  
Db 19529 EPPEIDLDSMRKLVIVRAGCPIRLFAIVRGRPAKVTWRKVGIDNVVKGQVDLDTMA 19588  
QY 190 ----- 189  
Db 19589 FLVIPNSTRDDSGKYSLTLVNPAGEKAVFVNVVRVLDTPGPGVSDLVKSDVTKTSCHVSWAP 19648  
QY 190 ----- 189  
Db 19649 PENDGGSQVTHYIVEKREADRKTWSTVTPVVKTSFHVTVNLVPGNEYFRVTAVNEYGPG 19708  
QY 190 ----- 189  
Db 19709 VPTDVPKPVLASDPLSEPPPRKLEVTMTKNSATLAWLPLRDGGAKIDGYITSYREE 19768  
QY 190 ----- 189  
Db 19769 QPADRWTEYSVVKDLSLVVTGLKEGKKYKFRVAARNAVGVSLEPREAEGVYEAKEQLLPK 19828  
QY 190 ----- 189  
Db 19829 ILMPEQITIKAGKRLIEAHVYGRPHPTCKWKKEDEVVTSLSHLAVHKADSSSILIIKDV 19888  
QY 190 ----- 189  
Db 19889 TRKDSGYISLTAENSSGTDQKIKVVVMDAPGPPPPFIDISDADACSLSWHIPLEDGG 19948  
QY 190 ----- 189  
Db 19949 SNITNYIVEKCDVSRGDWVTALASVTKTSRCKLIPGOEYIFRVRAENRFGISEPLTSP 20008  
QY 190 ----- 189  
Db 20009 KWAQFPFGVPSEPKNARVTKVNKDCIFVAWDRPDSGSGSPIGYLIERKERNILLSWKA 20068  
QY 190 ----- 189

Db 20069 NDTLVRSTEYPCAGLVEGLEYSFRIYALNKNAGSSPPSKPTEYVTAAMPVDPGKPEVIDV 20128  
QY 190 ----- 189  
Db 20129 TKSTVSLIWARPKHDGSKIIGYFVEACKLPGDKWVRCNTAPHQIPOEYATATGLEEKAQ 20188  
QY 190 ----- 189  
Db 20189 YOFRAIARTAVNISPPSESDPVTILAENVPRIDLSVAMKSLLTVKAGTNVCLDATVFG 20248  
QY 190 ----- 189  
Db 20249 KPMPVSWKKDGTLLKPAEGIKWAMQORNLCITLFSVNRKDSGTYTITAENSSGSKSATI 20308  
QY 190 ----- 189  
Db 20309 KLVLDKPGPASVKINKMYSDRAMLSWEPLEDGSGSEITNYIVDKRETSRPNWAQVSAT 20368  
QY 190 ----- 189  
Db 20369 VPITSCSVEKLJEGHEYOFRICAENKYGDPVFTEPAIAKNPYDPPGRCDPPVISNITK 20428  
QY 190 ----- 189  
Db 20429 DHMTVSWKPPADGGSPITGYLLEKRETOAVNWTKNRKPILIERTLKATGLOEGTEYEPR 20488  
QY 190 ----- 189  
Db 20489 VTAINKAGPGKSDAKAAYARDPOYPAPPAPFKYDTRSSVLSWCKPAYDGGSPIL 20548  
QY 190 ----- 189  
Db 20549 GYLVEYKRADSDNWVRNLPQNLOKTRFEVTGLMEDTOYQFRVYAVNKIGYSDPSVDPK 20608  
QY 190 -YPSDIIV----- 196  
Db 20609 HYPKDLIPPEGELDADLRKTLILRAGVTMRLYVPVKGRPPKTIWSPKNVNLDRIGLD 20668  
QY 197 ----- 196  
Db 20669 IKSTDFTLRCENVNKYDAGYILTLNESCCKKEYTIWVKYLDTPGPPVNTVKEISKD 20728  
QY 197 ----- 196  
Db 20729 SAYVTWEPPIIDGSPFIINYVQKRAERKSWSTVTTECKTSFRVANLEEGSYFFRVF 20788  
QY 197 ----- 196  
Db 20789 AENEYIGDPGETROAVKASQTPGPVVDLKVRSVSKSCSIGWKKPKHSDGGSRIIGYVVD 20848  
QY 197 ----- 196  
Db 20849 FLTEENKQRMVMSLSQYSAKDLTEGKEYTFRVSAENENGEGTPTSEITVVARDVVPD 20908  
QY 197 ----- 196  
Db 20909 LDLKGLPDLCYLAKENSFRLLKPIKGPAPSVSNKKGEDPLATDTRVSVSSAVNTLLI 20968  
QY 197 ----- 196  
Db 20969 VYDQKSDAGYTTITLKNVAGTKEGTISIKVVVKPGIPTGPIKFDVTAEMTLKWAPPK 21028  
QY 197 ----- 196  
Db 21029 DDGSEITNYILEKRDVNNKWWTCASAVQKTTFRVTRLHEGMEYTFRVSANENKYGVEG 21088  
QY 197 ----- 196  
Db 21089 LKSEPIVARHPDVPDAPPPNIVDVHDSVSLTWTDPKKTGSPITGVHLEFKERNILL 21148  
QY 197 ----- 196  
Db 21149 WKRANKTPIRMKDFKVTGLTEGLEYEFRVMAINLAGVKPSLPSEPVVALDPIDPPGKE 21208

QY 197 ----- 196  
Db 21209 VINITRNSVTLIWTEPKYDGGHKLTCYIVEKRDLPKSKMKNHVNVPCEAFTVTLVEG 21268  
QY 197 ----- 196  
Db 21269 GKYEFRIAKNTAGIASPSESTETIICKDEYEAPTIVLDPTIKOGLTIKAGDTIVLNAI 21328  
QY 197 ----- 196  
Db 21329 SILGRPLPKSSWSKAGKDIRPSDITQITSTPTSSMLTIKYATRCKDAGEYITITATNPFGTK 21388  
QY 197 ----- 196  
Db 21389 VEHVKVTLVDVFGPGPVEISNVSAAKATLTWTPPLEDGGSPIKSYILEKRETSRLWTV 21448  
QY 197 ----- 196  
Db 21449 VSEDIQSCRHVATKLIQONEYIFRVSANVHYKGEPVQSEPVKMVDRFGPPGPEKPEVS 21508  
QY 197 ----- 196  
Db 21509 NVTKNTATVSWKRPVDDGGSEITGYHVERREKSLRWVRAIKTPVSDLRCKVTGLQEGST 21568  
QY 197 ----- 196  
Db 21569 YEFVSAENRAGIGPPSEASDSVLMKDAAYPPGPPSPNHVTDTKKSASLAWGKPHYDGG 21628  
QY 197 ----- 196  
Db 21629 LEITGYVVEHOKVGEAWIKDITGTALRITQFVVPDLOTKERYNFRISAINDAGVGEPAV 21688  
QY 197 ----- 196  
Db 21689 IPDVEIVEREMADPELDAELRRTLWVRAGLSIRIFVPIKGRPAPEVTTKONILKNRA 21748  
QY 197 ----- 196  
Db 21749 NIENTESFTLLIPECNRYDTGKFVMTIENPAGKKSGFVNVRVLDTPGPNLNRPTDITK 21808  
QY 197 ----- 196  
Db 21809 DSVTLHWDPLIDGSRITNYIVEKREATRKSYSATTCHKCTYKVYVTLGSEGEYFFRV 21868  
QY 197 ----- 196  
Db 21869 MAENEYIGEPTETTEPVPKASEAPSPDLSNIMDITKSTVSLAWPKPKHGGSKITGYVI 21928  
QY 197 ----- 196  
Db 21929 EAQRKSGDQWTHITTVKGLECVVRNLTEGEYTFQVMAVNSAGRSAPRESRPVIVKEQTM 21988  
QY 197 ----- 196  
Db 21989 LPELDLRGIYQKLVIAKAGDNKIVEIPVLRPKPTVTWTKGDOILKQTVNFETTATST 22048  
QY 197 ----- 196  
Db 22049 ILNINECVRSDSGPPLFARNIVGEVGDVITIQVHDIPGPTGPIKFDVSVDFVTFSWD 22108  
QY 197 ----- 196  
Db 22109 PPENDGGVPISNVVVMRQTDSTTWELATTVIRITYKATRLTTGLEQYFRVKAQNRGV 22168  
QY 197 ----- 196  
Db 22169 GPGITSACIVANYPKVPGPPTQVAVTKDSMTISWHEPLSDGSPILGYHVERKERN 22228  
QY 197 ----- 196  
Db 22229 GILWOTVSKALVPGNIFKSSGLTDGIAVEFRVIAENMACKSKPSKPSEPMALDIPDPPG 22288

QY	197	-----	196
Db	22289	KPVLNITRHTVILKWKPEYTGFKITSYIVEKRDLPNGRWLKANFSNILENEFTVSGL	22348
QY	197	-----	196
Db	22349	TEDAAVEFRVIAKNAAGISPPSEPSDAITCRDDVEAPKIKVDVKFDTVILKAGEAFL	22408
QY	197	-----EWESNGO-----	203
Db	22409	EADVSGRPPPTMEWSKDGKELEGTAKLEIKIADFSTNLVKNKSTRDSDGAYTLTATNPGG	22468
QY	204	-----	203
Db	22469	FAKHIFNVKLDLRPGPEGLAVTEVTSEKCVLSWPPPLDDGGAKIDHYIVQKRETSRLA	22528
QY	204	-----	203
Db	22529	WTVASEVOVTKLVTKLKLKGNEYIFRMAVNKYGVGELESEPLAVNPYPGPPKPN	22588
QY	204	-----	203
Db	22589	EVTIITKDSMVWCWGHDPDSDGSGSEIINYIVERRDKAGORWIKCNKKTLDLRYKVSGLTE	22648
QY	204	-----	203
Db	22649	GHEVEPRIMAENAGISAPSTSPFYKACDTVFKPGPPGNPRVLDTSRSSISIANWKPIY	22708
QY	204	-----	203
Db	22709	DGSEITGYMVEIALPEDEWQIOTPPAGLKATSYTITGLTENQEKIRIYAMNSEGLGE	22768
QY	204	-----	203
Db	22769	PALVPGTPKAEDRMLPPEIELDADLRKVVTIRACCTLRFLVPIKGRPAPEVKWARDGES	22828
QY	204	-----	203
Db	22829	LDKASIESTSYTLLIIVGNVNRFDGKYILTVENSNGSKSAFVNVRYLDTGPPQDLVKY	22888
QY	204	-----	203
Db	22889	EVTKTSVTLTWDPPPLDGGSKIKNYIVEKRESTRKAYSTVATNCHKTSWKVDQLQEGCSY	22948
QY	204	-----	203
Db	22949	YFRVLAENEYIGILPRAETAESVKASERPLPPGKITLMDVTRNSVSLSWKPEHDGGSRL	23008
QY	204	-----	203
Db	23009	GYIVEMQTKGSKWATCATVKVTEATITGLIOGEYSFRVSAQNEKIGSDPRQLSVPVIA	23068
QY	204	-----	203
Db	23069	KDLVPPAFKLLFNTFTVLAGEDLKVDPVFIGRPTPAVTHKONVPLKQTRVNAESTEN	23128
QY	204	-----	203
Db	23129	NSLLTIKDACREDVGHVVKLTNSAGEAIEITLNVIVLDKPGPTGPVKMDEVTAADTSL	23188
QY	204	-----	203
Db	23189	WGPPKYDGGSSINNYIVEKRDTSTTTWQIVSATVARTTIKACRLKTGCEYQFRIAAENRY	23248
QY	204	-----	203
Db	23249	GKSTYLNSEPTVAQYKPKVPGPPTPVVTLSSRDSMEVQWNEPISDGGSRVIGYHLERKE	23308
QY	204	-----	203
Db	23309	RNSILWVKLNKTPIPQTKFTTGLBEGVEYEFVRVSAENIVGICKPSKVSECYVARDPCDP	23368
QY	204	-----	203

Db	23369	PGRPEALIVTRNSVTILOWKKPTYDGGSKITGYIVEKKEPELGRMKASFNTIIDTHEFTV	23428
QY	204	-----	203
Db	23429	GLVEDHRYEPRVIAARNAAGVSEFSESTGAITARDEVDPPRISMDPKYKDTIVVHAGESF	23488
QY	204	-----	203
Db	23489	KVDADIYGRPIPTIOWIKGDOELSNARLEIKSTDFATSLSDKAVRVDGNGVILKAKNV	23548
QY	204	-----	203
Db	23549	AGERSVTNVKVLDRPGPEGVPVVISGVTAEKCTLAWKPLQDGGSDIINYIVERRETSR	23608
QY	204	-----	203
Db	23609	LWTVVDANVQTLCKVKTKLLEGNEYTRIMAVNKGVGGEPLSEPVVAKNPFVVPDAPK	23668
QY	204	-----	203
Db	23669	APEVTTVTKDSMIVVWERPASDGGSEILGYVLEKRDKEGIRWTRCHKRLIGELRLRVGL	23728
QY	204	-----	203
Db	23729	IEHNDYFRVSAENAGLSEPPSPSAVQKACDPIYKPGPPNNPKVIDITRSSVFLSWSKP	23788
QY	204	-----	203
Db	23789	IYDGCIEIOGYIVEKCDVSYGEMTCTPTGTINKNTNIEVEKLEKHEYNFRICAINKAGV	23848
QY	204	-----	203
Db	23849	GEHADVPPIIVEEKELEAPDIDLDELRLKIINIRAGSLRLFPVPIKGRPTPEVKWGVDC	23908
QY	204	-----	203
Db	23909	EIRDAALIDVTSFTSLVDNVNRDYGKYTLTLENSSGTKSAFVTVRVLDTPSPVNLK	23968
QY	204	-----	203
Db	23969	VTEITKDSVSIWEPPPLDGGSKIKNYIVEKREATRKSAAVVTNCHKNKSWKIDQLOEGC	24028
QY	204	-----	203
Db	24029	SYFVRTAENEYIGILPAQTADPIKVAEVPQPPCKITVDDVTNRNSVSLSWTKPEHDGSK	24088
QY	204	-----	203
Db	24089	IIQYIVEMQAKHSEKSEKARVKSLOAVITNLTOGEEYLFVRVAVNEKGRSDPRSLAVPI	24148
QY	204	-----	203
Db	24149	VAKDLVTEPDVKPAFSSYSVQVGODLKIEVPIISGRPKPTITWTKDGLPLKQTRINVTDS	24208
QY	204	-----	203
Db	24209	LDLTLSIKETHKDDGGQYGITVANVVGOKTASIEIVTLDPDPKPGPKVDFDVSARSIT	24268
QY	204	-----	203
Db	24269	LSWNPPLYTGGCQITNYIVQKRDTTWTVDWSATVARTTLKVTKLKTGTGYQFRIFAEN	24328
QY	204	-----	203
Db	24329	RYQCSFALES DPIVAQYYPKEPGPGTFFATAISKDSMWIQWHEPVNNGSPVIGYHLER	24388
QY	204	-----	203
Db	24389	KERNILWTKVNTIITHDTQFKAQNLKEEGIEYEFVRVAENIVGVKASKSECYVARDPC	24448
QY	204	-----	203

Db	24449	DPPGTPPIWKRNEITLQWTKPVYDGGSMITGYIVEKRDLPDGRWMAKSFNTVETQFT	24508
QY	204	-----	203
Db	24509	VSGLTEDQRYEYFRVIAKNAAGALSPSDSTGPITAKDEVLPRIISMDPKFRTDITVIVNAGE	24568
QY	204	-----	203
Db	24569	TFRLADVHGKPLPTIEMLRGDKIEBESARCEIKNTDFKALLIVKDAIRIDGGQYILRAS	24628
QY	204	-----	203
Db	24629	NVAGSKSPVNVKVLDRPPEGPVQVGTGTSEKCSLTWPPQLDGGSDISHYVVEKRET	24688
QY	204	-----	203
Db	24689	SRLAWTVASEVVTNSLKVTKLEGNEYFRIMAVNKNYGVGEPLSAPVLMKNPFVLPQP	24748
QY	204	-----	203
Db	24749	PKSLEVTNIAKDSMTVCMNRPDGSGSEIIGYIVEKRDRSGIRWIKCNKRRTDRLRLVT	24808
QY	204	-----	203
Db	24809	GLTEDHEYFRVSAENAGVGFSPATVYKACDPVKFPPTNAHIVDTTKNSITLAWG	24868
QY	204	-----	203
Db	24869	KPIYDGGSEILGVVVEICKADEEHWQIVTPQGLRVTRFEISKLETHEQYKIRVICALNKV	24928
QY	204	-----	203
Db	24929	GLGEATSPGTVKPEDKLEAPELDLDSELKGIIVVRAGSARIHIPFKGRPTPEITWSRE	24988
QY	204	-----PEN 206	
Db	24989	EGEFTDKVQIEGVNVTQLSIDNCRNDAGKYILKLENSGSKSAFVTVKVLDTPGPPQN	25048
QY	207	-----	206
Db	25049	LAVKEVRKDSAFVWEPIIDGAKVKNYVIDKRESTRKAYANVSKSKTSFKVENLTE	25108
QY	207	-----	206
Db	25109	GAIYYFRVMAENEFVGVDPVETVDAVKAAEPSPGKVTLTDVQSATSASLMWKEPHDGG	25168
QY	207	-----	206
Db	25169	SRVLGVVEMQPKGTEKWSIVAESKVCNAVVTGLSSGOEYQFRVKAYNEKGKSDPRVLGV	25228
QY	207	-----	206
Db	25229	PVIADLTIQPSKLKLPFNITYIOAGEDLKIEIPVIGRPRPNISWVKDGEPLKQTTVRNVE	25288
QY	207	-----	206
Db	25289	ETATSTVLHIKEGNKDDFGKYVTATNSAGTATENLSVIVLEKPGPPGVPRFDEVSADF	25348
QY	207	-----	206
Db	25349	VVISWEPPAYTGGCQISNYIVEKRDTTTTTHWMSATVARTTIKITKLKTGTEYQFRIFA	25408
QY	207	-----	206
Db	25409	ENRYGKSAPLDSKAVIVQYPFKEPGPGTPTFTSISKDQMLVQWHEPVNDGGTKIIGYHL	25468
QY	207	-----	206
Db	25469	EQEKNSILWLNKLTPIQDTRFKTTGLDGLGYEYBKFVSAENIVGIGKPKYSECFVARD	25528
QY	207	-----	206
Db	25529	PCDPPGRPAIVITRNNVTLKWKPAYDGGSKITGYIVEKKDLPDGRWMAKSFNTVLETE	25588
QY	207	-----	206
Db	25589	FTVSGLVEDQRYEYFRVIARNAAGNFSEPSDSSGATARDEIDAPNASLDPKVKDVLVHA	25648
QY	207	-----	206
Db	25649	GETFVLEADIRGKPIPDVVVMSKDGKELEETAARMEIKSTIOKTTLLVVKDCIRTDGGQYIL	25708
QY	207	-----	206
Db	25709	KLSNVGGTKSIDITVVKVLDRPPEGPPLKVTGVTAEKCYLAWNPPLQDGGANISHYIEK	25768
QY	207	-----NYKTT-----	211
Db	25769	RETSRLSWTQVSTEVAQALNYKVTLLPGNEYIFRVMVKNKYGIGEPLESGPVTACNPYPK	25828
QY	212	-----	211
Db	25829	PGPPSTPEVSATIKDSMVVTWARPVDDGTEIEGYILEKRDKEGVWTKCNKKTLLTDLRL	25888
QY	212	-----	211
Db	25889	RVTGLTEGHSYEFRAENAAGVGEPSFVSFYRACDALYPPGPPSPNKVTDTSRSSVSL	25948
QY	212	-----	211
Db	25949	AWSKPIYDGGAPVKGVVVEVKEAADEWTTCTPTGLOGKQFTVTKLKENTENFRICAI	26008
QY	212	-----	211
Db	26009	NSEGVGEPATLPGSVVAQERIEPPEIELDADLRKVVVLRASATLRLFTVTKGRPEPEVKW	26068
QY	212	-----	211
Db	26069	EKAEGILTDRAOIEVTSFTMLVIDNTRFDSGRYNLTLENNSGSKTAFVNVVRVLDSPSA	26128
QY	212	-----	211
Db	26129	PVNLTIREVKKDSVTLISWEPPLIDGGAKITNYIVEKRETRKAYATITNCTKTTFRIEN	26188
QY	212	-----	211
Db	26189	LOEGCSYFRVLASNEYIGLPAETTEPVKVSEPPPLPGRVTLVDVTRNTATIKWEKPES	26248
QY	212	-----	211
Db	26249	DGGSKITGVVEMQPKGSEKWSCTQVKTLEATISGLTAGEEYVFRVAANKEGRSDPRQ	26308
QY	212	-----	211
Db	26309	LGVPIARDIEIKPSVELPFFHTFNVKAREQLKIDVPFKGRPOATVNRKDGOTLKETTRV	26368
QY	212	-----	211
Db	26369	NVSSSKTVTSLISKEASKEDVGTYELCVNSAGSITVPTIIVLDRPGRPPIRIDEVSC	26428
QY	212	-----	211
Db	26429	DSITISWNPPEYDGCQISNYIVEKKTSTTWHIVSOAVARTSIKIVRLTTGSEYQFRV	26488
QY	212	-----	211
Db	26489	CAENRYGKSSYSESAVVAEYFPFPPTGPKVVHATKSTMLVTWQVPVNDGGSRVIGY	26548
QY	212	-----	211
Db	26549	HLEYKERSILWSKANKILIAIDTOMKVSGLDEGLMYEYRVAENIAGIKCSKSCPEVPA	26608
QY	212	-----	211
Db	26609	RDPDPPGQPEVTNITRKSIVSLKWSKPHYDGGAKITGYIVERRELPGDRLWLCNVTNIQE	26668

Qy 212 ----- 211  
Db 26669 TYFEVTELTQRYEERFARNADSVSESESTGPIIVKDDVEPPRVMQDKFRDVIW 26728  
Qy 212 ----- 211  
Db 26729 KAGEVLKINADIAGRPLPVISWAKDGEIERARTEIISTDNHPTLLTVKDCIRRDQTQYV 26788  
Qy 212 ----- 211  
Db 26789 LTLKNAVAGTRSAVAVCKVLDKPGPPAGPLEINGLTAEKCSLSWGRPOEDGGADIDYIVE 26848  
Qy 212 ----- 211  
Db 26849 KRETSHLAWTICEGELQMTCKVTKLLKGNEYIFRVTGVNKGYGEPLESVAIKALDPFT 26908  
Qy 212 ----- 211  
Db 26909 VPSPTSLTITSVTKESMTLWSPESDGGSEISGYIIEREKNSLRVRVKKPVVDLR 26968  
Qy 212 ----- 211  
Db 26969 VKSTGLREGCEYEVYVNAAGLSLPSETSPILRAEDPVFLPSPSKPIVDSGKTTIT 27028  
Qy 212 ----- 211  
Db 27029 IAWVKPLFDGGAPITGYTVYKKSDDTDWKTISIQLRGTEYTIISGLTTGAEYVFRKSVN 27088  
Qy 212 ----- 211  
Db 27089 KVGASDPSDDPOIAKEREPEPLFIDSEMRKTLIVKAGASFTMTVFRGRPVPNLWS 27148  
Qy 212 ----- 211  
Db 27149 KPDTDLRAVVDTSRTSLTIENANRDSGKTYLTIONVLSASLTLLVKVLDTPGPP 27208  
Qy 212 ----- 211  
Db 27209 TNIQTQDVTKESAVLSWDVPENDGGAPVKNYHIEKREASKAWSVTNNCNRLSYKVTNL 27268  
Qy 212 ----- 211  
Db 27269 QEGAIYFVRVSGENEFVGIPAETKEGVKITEKPSPEKLGVTISKDSVSLTWLKPETHD 27328  
Qy 212 ----- 211  
Db 27329 GGSRIYHVVEALEKGQKNWKCAVAKSTHHVVSGLRSENSEYFFRVFAENQAGLSDPREL 27388  
Qy 212 ----- 211  
Db 27389 LLPVLKEQLEPPEIDMKNFPHTVYVRAGSNLKVDPISGKPLPKVTLSRDGVPLKATM 27448  
Qy 212 ----- 211  
Db 27449 RFNTEITAENLTINLKESVTADAGRYEITTAANSSTGTTKAFINIVLDRPGPTGPVVISD 27508  
Qy 212 ----- 211  
Db 27509 ITEESVILKWEPPKYDGSQVTNVLKRETSTAVWTEVSATVARTVMKVMKLTITGEEYQ 27568  
Qy 212 ----- 211  
Db 27569 FRIKAENRFGISDHDSACVTKLPTYTTPGPPSTPWNTVNTRESITVGMHEPVSGGSVAV 27628  
Qy 212 ----- 211  
Db 27629 VGYHLEMKDRNSILWQANKLIVIRTHFKVTTISAGLIYEFVRYAENAAAGVKPSHPSEP 27688  
Qy 212 ----- 211  
Db 27689 VLADACEPPRNVRITDISKNSVLSWQOPAFDGGSKITGYIVERRDLDPGRWTKASFTN 27748  
Qy 212 ----- 211

Db 27749 VTETQFIISGLTONSQYEERFARNAVGSIISPSEVVGPITCIDSYGGVIDLPLEYTEV 27808  
Qy 212 ----- 211  
Db 27809 VKYRAGTSVKLRAGISGKPAPTIIEWYKDDKELOTNALVCVENTTDLASILIKDADRLNSG 27868  
Qy 212 ----- 211  
Db 27869 CYELKLRNMGASATIRVOILDKPGPPGPIEFKTVTAEKITILLWRPPADDGGAKITHY 27928  
Qy 212 ----- 211  
Db 27929 IVEKRETSRVVWSMVSEHLEBICIITTTIKIGNEYIFRVRAVNYKIGEPLESDSVVAKN 27988  
Qy 212 ----- 211  
Db 27989 AFVTPGPGPIPEVTKITKNSMTVWWSRPIADGGSDISGYFLEKRDKKSLGFKVLKETIR 28048  
Qy 212 ----- 211  
Db 28049 DTRQKVTGLTENSQYQYRVCAVNAAGOGPFSEPFYKAAADPIDPPGPAKIRIADSTKS 28108  
Qy 212 ----- 211  
Db 28109 SITLGSKPVYDGGSAVTGYVVEIROGEEETVSTKGEVRTTEYVYVSNLPGVNYFYR 28168  
Qy 212 ----- 211  
Db 28169 VSAVNCAGOGPIEMNEFVQAKDILEAPEIDLVALRTSVIAKAGEDVQVLIIPFKGRPPP 28228  
Qy 212 ----- 211  
Db 28229 TVTWKDEKNLGDARYSIENTDSSLLTIPQVTRNDTGKYLTIENGCVGEPKSTVSVK 28288  
Qy 212 ----- 211  
Db 28289 VLDTPAACQKLQVKHVRGTVTLTLLWDPPLIDGSPPIINVIYIEKRDATKRTWSVVSHKCSS 28348  
Qy 212 ----- 211  
Db 28349 TSFKLIDLSEKTPFFFRVLAENEIGIEGPCETTEPVRKAAEVPAPIRDLMSKDKTSVIL 28408  
Qy 212 ----- 211  
Db 28409 SWTKPDPDGGSVITEYVVERKKGEGTWSHAGISKTCEIEVSQLEQSVLEFRVFAKNEK 28468  
Qy 212 ----- 211  
Db 28469 GLSDPVTIGPITVKELIITPEVDLSDIPGAQVTVRIGHNVHLELPYKGPSPISWLKDG 28528  
Qy 212 ----- 211  
Db 28529 LPLKESEFVRFSKTEKNKTLISIKNAKKEHGKGYTVILDNAVCRIVATPITVITLGPSPKPK 28588  
Qy 212 ----- 211  
Db 28589 GPIRFDELKADSVILSWDVPEDNGGGEITCYISIEKRETSQTNWKMVCSSVARTTFKVPNL 28648  
Qy 212 ----- 211  
Db 28649 VKDAEYQFRVRAENRYGVSOPLVSSIIVAKHQFRIPGPGKPVYINVTSDGMSLTWDAPV 28708  
Qy 212 ----- 211  
Db 28709 YDGGSEVTGFHVEKKEKERNILWQKNTSPISGREYRATGLVEGLDYQFRVYAENSAGLSS 28768  
Qy 212 ----- 211  
Db 28769 PSDPSKFTLAVSPVDPGTPDYIDVTRITILKNWPNPLDGGSKIVGYSIEKROGNERWV 28828  
Qy 212 ----- 211  
||:

Db 28829 RCNFTDVSECQYTVTGLSPGDRYEFRIARNNAVGTISPPSQSGIIIMTRDENVPPIVEFG 28888  
QY 215 ----- 214  
Db 28889 PEYFDGLIISGBSLRIKALVQRPVPRVTFWKGVEIEKRMNMEITDVLGSTSLEVRDA 28948  
QY 215 ----- 214  
Db 28949 TRDRHGVTVEAKNAGSAAEIKVKVQDTPGKVVGPIRFTNITGKMTLWMDAPLNDGC 29008  
QY 215 -----LSD 218  
| | | |  
Db 29009 APITHYIEKRETSRLAWALIEDKCAQSYTAIKLINGNEYQFRVSAVNKFGVGRPLSD 29068  
QY 219 ----- 218  
Db 29069 PVVAQIQYVDPDAGIPESNITGNSITLTWARPESDGGSEIQOYLIERREKKSTRWVKV 29128  
QY 219 ----- 218  
Db 29129 ISKRPISETRFKVTGLTEGNEYEFHVMAENAGVGPASGISRLIKCREPVNPPGPTVVK 29188  
QY 219 ----- 218  
Db 29189 VTDTSKTTSLEWSKPVFDGMEIIGYIIMCKADIGDWHKVNAEACVKTRYTVDLQAG 29248  
QY 219 ----- 218  
Db 29249 EYKFRVSAINGAGKDSCEVTGTIKAVDRLTAPELDIDANFKQTHVVRAGASIRLFIAY 29308  
QY 219 ----- 218  
Db 29309 QGRPTPTAVMSKPSNLSLRADIIHTDSEFTLTVCNCRNDAGKYTLTVENNSGSKSIF 29368  
QY 219 ----- 218  
Db 29369 TVKVLTPGPPGPIPKDVTGRSATLMDAPLLDGGARIHHYVVEKREASRSQVISEK 29428  
QY 219 ----- 218  
Db 29429 CTRQIFKVNDAEGVPYFRVSAVNEYGVGEPEYEMPEPIVATEQPAPPRLDVDTSKS 29488  
QY 219 ----- 218  
Db 29489 AVLAWLKDGGSRITGYLLEMROKGSDFWEAGHTKOLTFTVERLVEKTEYEFVRKAK 29548  
QY 219 ----- 218  
Db 29549 NDAGYSEPREAFSSVIIKEPQIEPTADLTGINTQLITKAGSPFTIDVPISGRPAKVTW 29608  
QY 219 -----GSFFL----- 223  
| | |  
Db 29609 KLEEMRLKETDRVSITTKDRITTLTVKDSMRDGSGRYFLTLENTAGVKTFSTVVVIGRP 29668  
QY 224 ----- 223  
Db 29669 GPTGPIEVSSVSAESCVLWSGEPKDGGEITINXIVKRESGTTAWQLVNSSVRKTOIK 29728  
QY 224 ----- 223  
Db 29729 VTHLTXYMEYSFRVSENREFGVSKPLESAPIIAEHFPVPPSAPTRPEVHVHSANAMSINW 29788  
QY 224 ----- 223  
Db 29789 EEPYHDGSKIIGYWEKERNITLWVKENKVPCLCNYKVTGLVEGLEYPQRTVALNAA 29848  
QY 224 ----- 223  
Db 29849 GYSKASEASRPIMQONPVDAPRPEVTDVTRSTVSLINSAPAYDGGKVGVYIERKKPVS 29908  
QY 224 ----- 223  
Db 29909 EVGDGRWLKCNITYISDNFTVTALSEGDTYEFVRVLAKNAAGVISKGSESTGPTCRDEY 29968

QY 224 ----- 223  
Db 29969 APPKAELDARLHGLDVTIRAGSDLVLDAAVGGKPEPKIITWTKGDKELDCEKVSLOYTK 30028  
QY 224 ----- 223  
Db 30029 RATAVIKFCDRSDSGKYTLTVKNASGKAVSVNVKVLDSGPCGKLTVSRVTQEKCTLAW 30088  
QY 224 ----- 223  
Db 30089 SLPQEDGGAETHYIVERRETSRLNMVIVEGECPTLSYVVTRLIKNNEYIFRVRANKYG 30148  
QY 224 ----- 223  
Db 30149 PGVPVESEPIVARNSEFTIPSPGCIPEEVGTGKEHIIIIQWTKPESDGGNEISNLYVDKREK 30208  
QY 224 ----- 223  
Db 30209 KSLRWTRVKNKYVVDTRLKVTSLMEGCDYQFRVTAVNAAGNSEPSEASNFISCREPSYT 30268  
QY 224 ----- 223  
Db 30269 PGPPSAPRVVDTKHSISLAWTKPMYDGTDIVGVLEMQEKDQOWYRVHTNATIRNTE 30328  
QY 224 ----- 223  
Db 30329 FTVPDLKMCQKYSFRVAANVVKMSEYSESAIEIEPVERIEIPDLEADDLAKTVTIRAG 30388  
QY 224 -----YSKLTVDK----- 231  
| | |  
Db 30389 ASRLMWVSYGRPPPVITWSKOGIDLASRAIIDTTESYSLIIVDKVNRDYDAGKYTAEAN 30448  
QY 232 ----- 231  
Db 30449 QSGKKSATVLVKVYDTPGPCPSVKVKEVSRDSVTITWEIPTIDGGAPVNNYIVERREAM 30508  
QY 232 ----- 231  
Db 30509 RAFKVTWTKCKTLYRISGLVEGTMYFRVLPENIYGIGPCETSDAVLSEVPLVPAKL 30568  
QY 232 ----- 231  
Db 30569 EYVDVTSTVTLAMEKPLYDGGSRLTGYVLEACKAGTERMMKVTLKPTVLEHTVTSNE 30628  
QY 232 ----- 231  
Db 30629 GEQYLFIRAQNEKGVSEPRETVTAVTVQDLRLVLPIDLSIMPQKTIHVPAGRPVELVIP 30688  
QY 232 ----- 231  
Db 30689 IAGRPPPAASWFFAGSKLRESERVTVETHKVKAKLTIRETTIRDTGEYTLKLNVTGTS 30748  
QY 232 ----- 231  
Db 30749 ETIKVIILDKPGPTGPIKIDEIDATISITISWPELDDGAPLSGYVVEQDRAHRPGMLP 30808  
QY 232 ----- 231  
Db 30809 VSESVTRSTFKTRLTGEGNEYFRVAATNRFGIGSYLQSEVIECRSSIRIPGPETLQIF 30868  
QY 232 ----- 231  
Db 30869 DVSRDGMTLTWYPEDDGGSQVTGYIVERKEVRADRVVRVKNVPVTMTTRYSTGLTEGLE 30928  
QY 232 ----- 231  
Db 30929 YEHRVTAINARGSGKPSRSPKPIVAMDPIAPPGKQONPRVTDTTTTSVSLAWSVPEDEGG 30988  
QY 232 ----- 231  
Db 30989 SKVTGYLEMQVDQHEWTKCNTTPTKIREYTLTHLPOGAEYFRVRLACNAGGPGPAEV 31048



```
QY 232 -----SRWQOG----- 237
Db 31049 PGTVKVTEMLEYDYELDERYQEGIFVRQGGVIRLTPIKGPFPICKWTKQODISKRA 31108
QY 238 ----- 237
Db 31109 MIATSEHTELVKEADRGSGYVDLVLENKCKKAVYIKRVIGSPNSPEGPFLYDDIQ 31168
QY 238 ----- 237
Db 31169 VRSVRVSRPPADGGADILGYILERREVPKAAWYITDSRVRGTSLVVKGLKENVEYHFR 31228
QY 238 ----- 237
Db 31229 VSAENQFISKLKSEBPVTKPLNPPPEPSPNPPELVDTKSSVLSWSRPKDDGSRV 31288
QY 238 ----- 237
Db 31289 TGYIYERKETSTDKWVRHNKTQTTTMYVTGLVPDAEYQFRIQAQNDVGLSETSPASEP 31348
QY 238 ----- 237
Db 31349 VVCKDPDKPSQPCLEILSISKDSYTLQWEKPECDDGKELLGYWVEYRQSGDSAMKKSXN 31408
QY 238 ----- 237
Db 31409 KERIKDKQFTIGLLEATEYEFVRFAENETGLSRPRRTAMSIKTLTSGEAPGIRKEMKD 31468
QY 238 ----- 237
Db 31469 VTTKLGBAALSQOIGRPLDIPKWRFGKELIQSRKRYKMSDGRTHLTVMTEEQDEG 31528
QY 238 ----- 237
Db 31529 VYTICATNEGEVETSSKLLQATPQHPGYPKLEKYGAVGSTLRHLHVMYIGRPVPAWT 31588
QY 238 -----NVF----- 240
Db 31589 WFHQKLLONSENITIENTEHTHLVMKNVQKTHAGKYKVQLSNVFGTVDAILDVEIQD 31648
QY 241 ----- 240
Db 31649 KDKPTGPIVIEALLKNSAVISMKPPADGGSWITNTYVVEKCEKAEGAEQWLVSSAISVT 31708
QY 241 ----- 240
Db 31709 TCRIVNLTENAGYFRVSAQNTFGISDPLEVSSVWIIKSPFEXPGAPKPTITAVTKDSC 31768
QY 241 ----- 240
Db 31769 VVAKPPASDGGAKIRNYLEKREKKONKWSVTTEIETVFSVKNLIEGLEVEFRVKC 31828
QY 241 ----- 240
Db 31829 ENLGESEWSEISEPIPKSDVPIQAPHFKELNLNRYQSNATLVCKVTGHPKPIVKW 31888
QY 241 ----- 240
Db 31889 YROGKEIIADGLKYRIQEFKGYHQLIIASVTDATVYQVRATNOGGSVSGTASLEVEV 31948
QY 241 ----- 240
Db 31949 PAKIHLPKTLEGMGAVHALRGEVVSIIKIPFSGKPDVITWQKQDLIDNNGHYQVIVTRS 32008
QY 241 ----- 240
Db 32009 FTSLVFPNGVERKDAGFYVCAKNRFGIDQKTVELDVADVPDPGRGVKVSVDVSRDSYNLT 32068
QY 241 ----- 240
Db 32069 WTEPASDGGSKITNYIVEKATTAERWLVRVQARETRTYVINLFGKTSYOFRIAENKFG 32128
QY 241 ----- 240
```

```
Db 32129 LSKPSEPSEPTITKEDKTRAMNYDEEYDEVTSMTKASHSTKELYKYEMIAEDLGRGE 32188
QY 241 ----- 240
Db 32189 FGIHRCVETSSKTYMAKFKVKGDQVLVKKEISILNTARHNILHLHESFESMEELY 32248
QY 241 ----- 240
Db 32249 MIFEFIGLDIFERINTSAFELNEREIVSVHQVCEALQFLHSHNIGHFDIRPENIIYQT 32308
QY 241 ----- 240
Db 32309 RRSSTIKIIEFGOARQLPGDNFRLFLTAPEYYAPEVHQHDVVSTATDMWSLGTLYVLL 32368
QY 241 ----- 240
Db 32369 SGINPLAETNQOIIENIMNAEYTFDEEAFKEISIEAMDVFDRLLVYKERSKRTASEALQ 32428
QY 241 ----- 240
Db 32429 HPWLKQKIERVSTKVIRTLAHRYYHTLIKLDLNMVVSAAIRISCGGAIRSQKGVSAKVK 32488
QY 241 ----- 240
Db 32489 VASIEIGPVSQIMHAVGEGGHVKYCKIENYDQSTQVTWYFGVROLENSEKYEITYED 32548
QY 241 ----- 240
Db 32549 GVAILYVKDITKLDGTYRCKVVDYGEDSYAELFKVGREVDYVYCRRTMKKIKRRTD 32608
QY 241 ----- 240
Db 32609 TMLLERPPETPLPLYNKTAIVGENVRFVGTITVHPPEHTWYKSGOKIKPGDNKKYTF 32668
QY 241 -----SC----- 242
Db 32669 ESDGLYQLTINSVTTDDAEYTVARNKYGEDSKAKLTVTLHPPTDSTLRPMFKRL 32728
QY 243 -----SYMHEAL----- 249
Db 32729 ANAEQBGQSVCFEIRVSGIPPTLKWEKDGQPLSGPNIEIIEHGLDYIALHIRDTLPE 32788
QY 250 ----- 249
Db 32789 DTGYRVATNTAGSTSCQAHQVERLYKKQEFKKEEHERHVOKIDKTLRMAEILSG 32848
QY 250 ----- 249
Db 32849 TESVPLTQVAKREALAAVLYKPAVSTKTVKGBFRLEIEBKERKLRMPYDVPEPRKYK 32908
QY 250 ----- 249
Db 32909 QTTTEEDQRIKQFVPMDSMKWYKKIRDQYEMPKLDRVQKRPKRIRLSRWEQFYVMPLP 32968
QY 250 ----- 249
Db 32969 RITDOYRPKWRIPKLSQDDLEIVRPARRRTPSPDYDFYRPRRSIGDISDEELLPLIDD 33028
QY 250 ----- 249
Db 33029 YLAMKRTTEERLLEEELELGSASPPSRPPHFELSSLRYSQPAHVKEETRKDFRYS 33088
QY 250 ----- 249
Db 33089 TYHIPTKAEASTSYAELRERHAQAAAYRQPKORIMAREDEELLRPVTTTQHLSYKSE 33148
QY 250 ----- 249
Db 33149 LDFMSKEKSKRSRRQREVTETEIEEYIEISKHAQRESSSSASRLRRRRSLSPYIE 33208
QY 250 ----- 249
```

```
Db 33209 LMRPVSELIRSPQAEYEDDTERSPPTTRPRSPSPVSSERSLSRFSARSARDIFS 33268
QY 250 -----HN----- 249
Db 33269 RYESKMAALKTKTKTSERKYEVLSSQPFLLDHAPRITLMRSHRVPCCGQNTREILNVQSKP 33328
QY 250 -----HN----- 251
Db 33329 TAENVKWHNGVLOESKIHNTSGVLTLEILDCHTDGTYRAVCTNYKGEASDYATL 33388
QY 252 ----- 251
Db 33389 DVTGGDYTTASQRRDEEVRSPFPELTREAYAVSFFKKTSEMEASSSVREVKSQMTET 33448
QY 252 ----- 251
Db 33449 RESLSYHSAEMKSALEKSLKSTTKIKITTLAARILTPRSTVTVYEGESARFS 33508
QY 252 ----- 251
Db 33509 CDTDGEPTVTWLRKGVLSTARRHQVTTTKYKSTFEISSVQASDEGNYSVVWENSECK 33568
QY 252 ----- 251
Db 33569 QEAFTLTQKARTEKAVTSPRVKSPPEPRVKSPPEVSKPRVKSPPEPSHPKAVSPTET 33628
QY 252 ----- 251
Db 33629 KPTPEKVQHLPSAPPKITQFLKAEASKEIAKLCVVESSLRAKEVTVYKDGKHLKEN 33688
QY 252 -----HY----- 253
Db 33689 GHFQFHSADGTYELKINNLNLTSDOGYVCEISGEGGTSKTLQPMGQAFKSIHEKVSKI 33748
QY 254 ----- 253
Db 33749 SETKKSQDKTTESTVTRKTEPRKAPETPISPKPVITGLQDTTVSSDVAKFAVKATGEPRP 33808
QY 254 ----- 253
Db 33809 TAIWTKDGAITQGGYKLSDEKGGFFLEIHKHTDTSGLYCTCYKNAGSVSSCKLTI 33868
QY 254 -----TQKSLSPK 264
Db 33869 KAIKDEAKGVSTOKTSETTPQK 33891

RESULT 14
ID Q10466 PRELIMINARY: PRT: 26926 AA.
AC Q10466;
DT 01-NOV-1996 (TREMBLrel. 01, Created)
DT 01-NOV-1996 (TREMBLrel. 01, Last sequence update)
DE 01-JUN-2002 (TREMBLrel. 21, Last annotation update)
DE Titin, heart isoform N2-B (EC 2.7.1.-) (Connectin).
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=HEART;
RX MEDLINE=96026330; PubMed=7569978;
RA Label S., Kolmer B.;
RA "Titins: giant proteins in charge of muscle ultrastructure and
RT elasticity";
FT Science 270:293-296(1995).
RN [2]
RP SEQUENCE OF 22277-25376 FROM N.A.
RX MEDLINE=92258380; PubMed=1582406;
RA Label S., Gautel M., Lakey A., Trinick J.;
RA "Towards a molecular understanding of titin.";
FT EMBO J. 11:1711-1716(1992).
RN [3]
RP SEQUENCE OF 1976-2014 FROM N.A.
RA Label S.;
RL Submitted (DEC-1994) to the EMBL/GenBank/DBJ databases.
RN [4]
RP CHARACTERIZATION.
RX MEDLINE=95331314; PubMed=7607248;
RA Gautel M., Castiglione-Morelli M.A., Pfuhl M., Motta A., Pastore A.;
RT "A calmodulin-binding sequence in the C-terminus of human cardiac
RL titin kinase.";
RL Eur. J. Biochem. 230:752-759(1995).
CC -|- FUNCTION: THIS GIANT MUSCLE PROTEIN MAY BE INVOLVED IN MUSCLE
CC ASSEMBLY AND IN MAINTAINING THE STRUCTURAL INTEGRITY OF
CC SARCOMERES. MAY HAVE PROTEIN KINASE ACTIVITY.
CC -|- ALTERNATIVE PRODUCTS: A NUMBER OF FORMS OF THIS PROTEIN ARE
CC PRODUCED BY ALTERNATIVE SPLICING WHICH DIFFER IN TISSUE
CC DISTRIBUTION. DIFFERENT SIZE TRANSCRIPTS MAY ALSO EXIST WITHIN ANY
CC ONE TISSUE. THE SEQUENCE SHOWN HERE IS THAT OF THE HEART ISOFORM
CC N2-B.
CC -|- TISSUE SPECIFICITY: MUSCLE-SPECIFIC.
CC -|- SIMILARITY: TO THE CATALYTIC DOMAINS OF OTHER SERINE/THREONINE
CC KINASES.
CC -|- SIMILARITY: BELONGS TO IMMUNOGLOBULIN SUPERFAMILY. CONTAINS 112
CC IMMUNOGLOBULIN C2-LIKE DOMAINS AND 132 FIBRONECTIN TYPE III-LIKE
CC DOMAINS.
CC EMBL: X64698; CAA45939.1; -.
CC EMBL: X83270; CAA58243.1; -.
CC EMBL: X64697; CAA45938.1; -.
CC EMBL: X90568; CAA62188.1; -.
CC EMBL: X64699; CAA45940.1; -.
CC HSP; P56276; ITLK.
CC InterPro: IPR000282; Cytok_receptor_2.
CC InterPro: IPR000719; Buk_pkinase.
CC InterPro: IPR000577; FGGY_kin.
CC InterPro: IPR003962; FnIII_repeat.
CC InterPro: IPR003961; FnIII.
CC InterPro: IPR003598; Ig_C2.
CC InterPro: IPR003600; Ig_like.
CC InterPro: IPR003006; Ig_MHC.
CC InterPro: IPR00129; Peptidase_S24.
CC InterPro: IPR002016; Peroxidase.
CC InterPro: IPR002290; Ser_thr_pkinase.
CC InterPro: IPR001245; Tyr_pkinase.
CC Pfam: PF00041; fn3; 132.
CC Pfam: PF00047; Ig; 91.
CC Pfam: PF00069; pkinase; 1.
CC PRINTS: PR00014; FNTYPEIII.
CC PRINTS: PR00726; LEXASERPTASE.
CC PRODOM: PD000001; Euk_pkinase; 1.
CC SMART: SM00060; FN3; 127.
CC SMART: SM00408; IGC2; 23.
CC SMART: SM00410; IGLike; 79.
CC SMART: SM00220; S_TKC; 1.
CC PROSITE: PS00933; FGGY_KINASES_1; UNKNOWN_1.
CC PROSITE: PS00038; HELIX_LOOP_HELIX; UNKNOWN_1.
CC PROSITE: PS00290; IG_MHC; UNKNOWN_1.
CC PROSITE: PS00435; PEROXIDASE_1; UNKNOWN_1.
CC PROSITE: PS50011; PROTEIN_KINASE_DOM; 1.
CC PROSITE: PS00109; PROTEIN_KINASE_TYR; 1.
CC Muscle protein; cytoskeleton; Structural protein; Calmodulin-binding;
KW Serine/threonine-protein kinase; Alternative splicing; Repeat;
KW Immunoglobulin domain; Phosphorylation.
FT DOMAIN 1370 1389 4 X 5 AA TANDEM REPEATS OF R-M-S-P-A.
FT DOMAIN 4429 4614 GLU/LYS/PRO/VAL-RICH.
FT DOMAIN 24731 25070 CATALYTIC.
FT DOMAIN 25030 25056 CALMODULIN-BINDING.
FT MOD_RES 1372 1372 PHOSPHORYLATION (BY PDPK) (POTENTIAL).
FT MOD_RES 1377 1377 PHOSPHORYLATION (BY PDPK) (POTENTIAL).
FT MOD_RES 1382 1382 PHOSPHORYLATION (BY PDPK) (POTENTIAL).
FT MOD_RES 1387 1387 PHOSPHORYLATION (BY PDPK) (POTENTIAL).
FT MOD_RES 26171 26171 PHOSPHORYLATION (POTENTIAL).
FT MOD_RES 26178 26178 PHOSPHORYLATION (POTENTIAL).
```

```
RN [3]
RP SEQUENCE OF 1976-2014 FROM N.A.
RA Label S.;
RL Submitted (DEC-1994) to the EMBL/GenBank/DBJ databases.
RN [4]
RP CHARACTERIZATION.
RX MEDLINE=95331314; PubMed=7607248;
RA Gautel M., Castiglione-Morelli M.A., Pfuhl M., Motta A., Pastore A.;
RT "A calmodulin-binding sequence in the C-terminus of human cardiac
RL titin kinase.";
RL Eur. J. Biochem. 230:752-759(1995).
CC -|- FUNCTION: THIS GIANT MUSCLE PROTEIN MAY BE INVOLVED IN MUSCLE
CC ASSEMBLY AND IN MAINTAINING THE STRUCTURAL INTEGRITY OF
CC SARCOMERES. MAY HAVE PROTEIN KINASE ACTIVITY.
CC -|- ALTERNATIVE PRODUCTS: A NUMBER OF FORMS OF THIS PROTEIN ARE
CC PRODUCED BY ALTERNATIVE SPLICING WHICH DIFFER IN TISSUE
CC DISTRIBUTION. DIFFERENT SIZE TRANSCRIPTS MAY ALSO EXIST WITHIN ANY
CC ONE TISSUE. THE SEQUENCE SHOWN HERE IS THAT OF THE HEART ISOFORM
CC N2-B.
CC -|- TISSUE SPECIFICITY: MUSCLE-SPECIFIC.
CC -|- SIMILARITY: TO THE CATALYTIC DOMAINS OF OTHER SERINE/THREONINE
CC KINASES.
CC -|- SIMILARITY: BELONGS TO IMMUNOGLOBULIN SUPERFAMILY. CONTAINS 112
CC IMMUNOGLOBULIN C2-LIKE DOMAINS AND 132 FIBRONECTIN TYPE III-LIKE
CC DOMAINS.
CC EMBL: X64698; CAA45939.1; -.
CC EMBL: X83270; CAA58243.1; -.
CC EMBL: X64697; CAA45938.1; -.
CC EMBL: X90568; CAA62188.1; -.
CC EMBL: X64699; CAA45940.1; -.
CC HSP; P56276; ITLK.
CC InterPro: IPR000282; Cytok_receptor_2.
CC InterPro: IPR000719; Buk_pkinase.
CC InterPro: IPR000577; FGGY_kin.
CC InterPro: IPR003962; FnIII_repeat.
CC InterPro: IPR003961; FnIII.
CC InterPro: IPR003598; Ig_C2.
CC InterPro: IPR003600; Ig_like.
CC InterPro: IPR003006; Ig_MHC.
CC InterPro: IPR00129; Peptidase_S24.
CC InterPro: IPR002016; Peroxidase.
CC InterPro: IPR002290; Ser_thr_pkinase.
CC InterPro: IPR001245; Tyr_pkinase.
CC Pfam: PF00041; fn3; 132.
CC Pfam: PF00047; Ig; 91.
CC Pfam: PF00069; pkinase; 1.
CC PRINTS: PR00014; FNTYPEIII.
CC PRINTS: PR00726; LEXASERPTASE.
CC PRODOM: PD000001; Euk_pkinase; 1.
CC SMART: SM00060; FN3; 127.
CC SMART: SM00408; IGC2; 23.
CC SMART: SM00410; IGLike; 79.
CC SMART: SM00220; S_TKC; 1.
CC PROSITE: PS00933; FGGY_KINASES_1; UNKNOWN_1.
CC PROSITE: PS00038; HELIX_LOOP_HELIX; UNKNOWN_1.
CC PROSITE: PS00290; IG_MHC; UNKNOWN_1.
CC PROSITE: PS00435; PEROXIDASE_1; UNKNOWN_1.
CC PROSITE: PS50011; PROTEIN_KINASE_DOM; 1.
CC PROSITE: PS00109; PROTEIN_KINASE_TYR; 1.
CC Muscle protein; cytoskeleton; Structural protein; Calmodulin-binding;
KW Serine/threonine-protein kinase; Alternative splicing; Repeat;
KW Immunoglobulin domain; Phosphorylation.
FT DOMAIN 1370 1389 4 X 5 AA TANDEM REPEATS OF R-M-S-P-A.
FT DOMAIN 4429 4614 GLU/LYS/PRO/VAL-RICH.
FT DOMAIN 24731 25070 CATALYTIC.
FT DOMAIN 25030 25056 CALMODULIN-BINDING.
FT MOD_RES 1372 1372 PHOSPHORYLATION (BY PDPK) (POTENTIAL).
FT MOD_RES 1377 1377 PHOSPHORYLATION (BY PDPK) (POTENTIAL).
FT MOD_RES 1382 1382 PHOSPHORYLATION (BY PDPK) (POTENTIAL).
FT MOD_RES 1387 1387 PHOSPHORYLATION (BY PDPK) (POTENTIAL).
FT MOD_RES 26171 26171 PHOSPHORYLATION (POTENTIAL).
FT MOD_RES 26178 26178 PHOSPHORYLATION (POTENTIAL).
```

FT MOD\_RES 26184 26184 PHOSPHORYLATION (POTENTIAL).  
FT MOD\_RES 26190 26190 PHOSPHORYLATION (POTENTIAL).  
FT CONFLICT 22277 22277 T -> P (IN REF. 2).  
FT CONFLICT 22449 22449 E -> G (IN REF. 2).  
FT CONFLICT 22454 22454 T -> Q (IN REF. 2).  
FT CONFLICT 23324 23324 S -> L (IN REF. 2).  
SQ SEQUENCE 26926 AA; 2993428 MW; D5EECD3254DF5523 CRC64;  
  
Query Match 33.9%; Score 485; DB 4; Length 26926;  
Best Local Similarity 0.7%; Pred. No. 0.0002;  
Matches 177; Conservative 44; Mismatches 42; Indels 23713; Gaps 55;  
  
QY 2 WLK----- 4  
|||  
Db 1543 WLKNSDIIIVPHKPIRIBGTGKAALKIDSTVSQDSAWYATATNAKAGRODTTRCKVNYE 1602  
  
QY 5 -----AFYDK----- 9  
|||  
Db 1603 VEFAPPEPERKLIIPRGTYRAKEIAAPELEPLHLRYGQEWEEGLDYDKEKQKPFPPKK 1662  
  
QY 10 ----- 9  
Db 1663 LLSRLRKRGPAHFECRLTPISDPTMVVWELHDGPLEANRLRMINEFGYCSLDYGVA 1722  
  
QY 10 ----- 9  
Db 1723 SRDSGIITCRATNKYGTDHTSATLIVKDEKSLVEESQLPEGRKGLQRIEELERMAHEGAL 1782  
  
QY 10 ----- 9  
Db 1783 TGVTTDOKEKQPDIVLYPEPVRVLEGETAFRCRVGTGYPQPKVWYLNGLIRKSRFR 1842  
  
QY 10 ----- 9  
Db 1843 VRYDGIHLDIYDCKSYDTGEVKTVAENPEGVIEHKVLEIOQREDFRSVLRAPEPRPE 1902  
  
QY 10 ----- 9  
Db 1903 FHWHEPGKLOFEVQKVRDPVDTTETKEVVKLKRAERITHEKVPESSEELRSFKRRTEBG 1962  
  
QY 10 ----- 9  
Db 1963 YYEATAVELKSRKKDESYEELLRTKDELHWTKELTEEKKALAECKIITPTFKPKD 2022  
  
QY 10 ----- 9  
Db 2023 IELSPSMEAPKIFERIQSQTVCQSDAHPFRVRVVGKPDPECEWYKNGVKIERSDRIYWW 2082  
  
QY 10 ----- 9  
Db 2083 PEDNVCELVRDVTAEADSASIMVKAINIAGETSSHAFLVLQAKQLITFTQELQDVVAK 2142  
  
QY 10 ----- 9  
Db 2143 DTWATFECTSEPFVKVWKYKMGMEVHEGDKYRMHSDRKVHFLSLITIDTSDAEDYSCVL 2202  
  
QY 10 ----- 9  
Db 2203 VEDENVKTTAKLIVEGAVVEFKELQDIEVPESYSSELECIIVSPENIEGKYHNDVELKS 2262  
  
QY 10 ----- 9  
Db 2263 NGKYTTISRGRQNLTVKDVTKEDQGEYSFVIDGKTKTCKLKMKPRTAILQLGLSDQKVC 2322  
  
QY 10 ----- 9  
Db 2323 EGDIVOLEVKVLSVESVEGWMKQGEVQPSDRVHIVIDKQSHMLLIEDMTKEDAGNYSPT 2382  
  
QY 10 ----- 9  
Db 2383 IPALGLSTGRSVSVSDVITPLKDVNVIEGTKAVLECKVSPDVTSVKMWLNDEQIKPD 2442  
  
QY 10 ----- 9

Db 2443 DRVQAIKGTQKORLVINRTHASDEGPKLIVGRVETNCNLNLSVEKIKIIRGLDLCTETQ 2502  
QY 10 ----- 9  
Db 2503 NVFVEVELSHSGIDVLWNFKDKRKPSKYKIEAHGKIYKLVNLMMKDDCKYTFYAGE 2562  
QY 10 ----- 9  
Db 2563 NMTSGKLIVAGGAIKPLTDQTVAESQEAFFEVEANPDPSKGWLRDGHKLPITNNIRSE 2622  
QY 10 ----- 9  
Db 2623 SDGHRRLIIAATKLDIGEYTYKVATSKTSAKLVEAVKIKKTLKNLTVTQTQDAVFTV 2682  
QY 10 ----- 9  
Db 2683 ELTHPNVKGVOVKNKGVVLESNEYAISVKGITYSLRIKNCIAIVDESUYGFRGLGASA 2742  
QY 10 ----- 9  
Db 2743 RLHVETVKLIKKPKDVTALENATVAFEVSVDHTVPVKWFKHSVEIKPSDKHRLYSERV 2802  
QY 10 ----- 9  
Db 2803 HKLMLQNISPDAGEYTAVVGOLECKAKLFVETLHITKTMKNIEVPETKTASFECEVSHF 2862  
QY 10 ----- 9  
Db 2863 NVPSMWLKNVGEIEMSEKFIIVQVKLHOLIIMNTSTEDSAEYTFVCGNDQVSATLVTVP 2922  
QY 10 ----- 9  
Db 2923 IMITSMKDINAEEKDTITFEVTNVYEGISYKWLKNGVEIKSTDQCMETKLTLSLNR 2982  
QY 10 ----- 9  
Db 2983 NVHFGDAADYTFVAGKATSTATLYVEARHIEFRKHIDIKVLEKKRAMFECEVSEPDITV 3042  
QY 10 ----- 9  
Db 3043 QMKDDQELQITDRIKIQEKYVHRLIIPSTRMSDAGKYTVVAGGNVSTAKLFEVGRDVR 3102  
QY 10 ----- 9  
Db 3103 IRSIKKEVQVIEKQRAVVEFEVNEDDVDAHVKDGLIEINQVOERHKYVVERRIHRMFTS 3162  
QY 10 ----- 9  
Db 3163 ETRQSDAGEYTFVAGNRSSVTLYVNAPEPPVQLQELQVTVQSGKPARFCAMISGRPOP 3222  
QY 10 ----- 9  
Db 3223 KISWYKEBOLLSTGFKCKFLHDGQEYTLILLIEAFPEDAAYTCEAKNDYGVATTASLSV 3282  
QY 10 ----- 9  
Db 3283 EYVEVSPDQEMPVYPAIITPLQDVTVTSEGQPARFCQCRVSGTDLKVSWSKDKKIKPSR 3342  
QY 10 ----- 9  
Db 3343 FFRMTQFEDTYQLEIAEAYPEDEGTYTFVANNAGVSVSTANLSLEAPESILHERIEQEI 3402  
QY 10 ----- 9  
Db 3403 EMEMKEFSSFLSAEEGLHSAELQLSKINETLELLESFVPTPKFDSKEGTGPIFIKE 3462  
QY 10 ----- 9  
Db 3463 VSNADISMGDVATLSVTVIGIPKPIQWFFNGVLLTPSADYKVFDPGDDHSLIILFTKLE 3522  
QY 10 ----- 9

Db 3523 DEGEYTCMASNDYKGTICSAYLKINSKGEKHDTETESAVALSKLEKLGPCPPHFLKELK 3582  
Qy 10 ----- 9  
Db 3583 PIRCAQGLPALFEYTVVGEPAFTVTFWKENQOLCTSVYTTIHNPNNGSGTFIVNDPQBED 3642  
Qy 10 ----- 9  
Db 3643 SGLYICKAENLMBSTCAAEALLVLEDDTMDTPTCKAKSTPEAPDFPQTPKGPAAVEAL 3702  
Qy 10 ----- 9  
Db 3703 DSEQEIATFVKDTILKAALITEENOQLSYEHIAKANELSSOLPLGAOELQILEQDKLTP 3762  
Qy 10 ----- 9  
Db 3763 ESTREFLCINGSIHFPKLEPSNQLQIVOSQKTSKEGILMPEEPETQAVLSDTEKIF 3822  
Qy 10 -----VAE----- 12  
Db 3823 PSAMSIQINSLTVEPLKTLAEPEGNYPOSSIIPPWMSYLTSAEVLSLKERTVSDTN 3882  
Qy 13 ----- 12  
Db 3883 REQVRTLOKQEAQASALILSQSLAEGHVESLQSPDWMISQVNYEPLVPSEHSCTEGGKILI 3942  
Qy 13 ----- 12  
Db 3943 ESANPLENAGODSAVRIEKGSLRFLPALEKEQVLLKEHSDNVVMPDQIIIESKREPA 4002  
Qy 13 ----- 12  
Db 4003 IKKVOEGRODLLSKESLLSGIPBEORNLKIQICRALQAASVASEQPLGFSEWLNRTEKV 4062  
Qy 13 ----- 12  
Db 4063 EVEAVNITQEPRHIMCWLVTSAKSVTEVTIIEDVDPQMANLKMELRDALCAIIEEI 4122  
Qy 13 ----- 12  
Db 4123 DILTAEGPRIOGAKTSLQEEMDSFGSQKVEPITEPEVESKYLISTEVSFYFNQSRVK 4182  
Qy 13 -----KLKEA----- 17  
Db 4183 YLDATPVTKGVASAVSDEKQDSLKPSEKESSESSEGTBEVATVKIQEAEGGLIKEDG 4242  
Qy 18 ----- 17  
Db 4243 PMIHTPLDVTVSEGDIVHLTTSITNAKEVNWYFENKLVPSDEKFKCLQDQNTYTLVIDK 4302  
Qy 18 ----- 17  
Db 4303 VNTEDHOGYYCEALNSGKTATSAKLTVVYKRAAPVIKRKTEPLEVALGHLAKFTCEIQS 4362  
Qy 18 -----FDWLKA----- 23  
Db 4363 APNVRFQWFKAGREIYESDKSIRRSKYISSLEILRTQVVDGCEYCTCKASNEYGSVCTA 4422  
Qy 24 ----- 23  
Db 4423 TLTVTVPGEKKVRKLLPERKPEPKEEVLKSVLKRREPEEPKVPKLEKVKKPAVPE 4482  
Qy 24 ----- 23  
Db 4483 PPPKPVVEEVPVTKRKRIPEPTKVPEIKPAIPLPAPEPKPKPEAEVKTIKPPVP 4542  
Qy 24 ----- 23  
Db 4543 EPTPIAAPVTVVVGKKAERAKAPKEAAKPGPIKGVPKTPSPIEAERRKLKRGSGEK 4602  
Qy 24 ----- 23  
Db 4603 PPDEAPTYQLKAVPLKFVKEIKDIILTESEFVGSSAIFECLVSPSTAITTMKDKGSNIR 4662

Qy 24 ----- 23  
Db 4663 ESPKHFRIADGKDRKLHIIDVQLSDAGEYTCVLRGNKKEKSTAKLVVEELPVRVKTLE 4722  
Qy 24 ----- 23  
Db 4723 BEVTVVRGQPLYLSCLENKERDVVRKDKIVVEKPGRIVPGVIGLMRALTIINDADDTDA 4782  
Qy 24 ----- 23  
Db 4783 GTTYVTVENANNLECCSCVKVVEVIRDWLVKPIRDQHVKPGTAIPACDTAKDPNKKWF 4842  
Qy 24 ----- 23  
Db 4843 KGYDEIPAEPNDKTEILRDGNHLYLKIKNAMPEDIAEVAVEIECKRYPAKLTLCEREVEL 4902  
Qy 24 ----- 23  
Db 4903 LKPIEDVTIYKESASFAEISEADIPQWMLKGELLRSPSTCEIKAEKGKRFLTLHKVK 4962  
Qy 24 ----- 23  
Db 4963 LDQAGEVLYQALNAITTAILTVEIELDFAVPLKDVTPPERQARFECVLTREANVWSK 5022  
Qy 24 ----- 23  
Db 5023 GPDIIKSSDKFDIADGKKHILVINDSQFDDDEGVYTAEEVGKTSARLFTVGTIRLKFMS 5082  
Qy 24 ----- 23  
Db 5083 LEDQTVKEGETATFVCELSHEKMHVWFKNDAKLHTSRTVLISSEKTHKLEMKEVTLDD 5142  
Qy 24 -----FYDKVAE----- 30  
Db 5143 ISQIKAOVKELSSAQKLVLEADPYFTVKLHDKTAVEKDEITLKCEVSKOVVWKFWDGE 5202  
Qy 31 ----- 30  
Db 5203 EIVSPKYSIKADGLRILKIKADLKGEYVDCGDKTKANVTVEARLIEVEKPLYG 5262  
Qy 31 ----- 30  
Db 5263 VEVFVGETAHEIELSEPDVHGOWKLGOPLTASPDCEIIEIDGKKHILHLCOLGTMTE 5322  
Qy 31 -----KLKE----- 34  
Db 5323 VSFQAAAKSAANLKVRELPLIFITPLSDVKVFEKDEAKFECEVSREPRTFRWLKGTQEI 5382  
Qy 35 -----AFMD----- 38  
Db 5383 TGDDEFELIKDGTKHSWIKSAAFEDAKYFMAEDKHTSGKLIIEGIRLKLFTPLKDV 5442  
Qy 39 ----- 38  
Db 5443 AKESAVFTVELSHDNIRVKFNQDLHTTRSVSMQDEKTHSITFKDLSIDDTSQIR 5502  
Qy 39 ----- 38  
Db 5503 VEAMGSSEAKLTVLEGDPYFTGKLQDYTGVEKDEVILQCEISKADAPVWFKDGEIKP 5562  
Qy 39 ----- 38  
Db 5563 SKNAVINTDGKRRMLILKALKSDIGQYTCDCGTDKTSGLDIEDREIKLVRPLHSVEVM 5622  
Qy 39 ----- 38  
Db 5623 ETETARFETISEDHIHANWKLKGEALLQTPDCEIKEEGKIHSILVHNCRLDQTTGGVDFQ 5682  
Qy 39 ----- 38  
Db 5683 AANVKSAAHLRVPRVTGILLRPLKDVTVTAGETATFDCELSYEDIPVEWYLKGGKLEPSD 5742

QY 39 ----- 38  
Db 5743 KVVPRSEGVHTLTLRDVKLEDAGEVQLTAKDFKTHANLFVKPEPVEFTKPLEDQTVBEG 5802  
QY 39 ----- KTHTC--- 43  
Db 5803 ATAVLECEVSRENAKVKFKNGTEILKSKYIEIVADGRVKLVIHDCPTEDIKTVTCDAK 5862  
QY 44 ----- 43  
Db 5863 DFKTSCNLNVPPHVEFLRPLTDLOVREKEMARFCELSRENKVKFKDGAIEIKKKY 5922  
QY 44 ----- 43  
Db 5923 DIISKGAVRILVINKCLLDBEAESCEVRTARTSGLTVLEEAFTKNLANIEVSETDT 5982  
QY 44 ----- 43  
Db 5983 IKLCEVSKPGAETWYKGBDEIIEFTGRYEILTEGRKKRILVIONAHLEDAAGNYNCRLPSS 6042  
QY 44 ----- 43  
Db 6043 RTDGVKVKHELAAEFISKPNLEILLEGEKAFFVCSISKESFPVQWKRDDKTILES GDKYD 6102  
QY 44 ----- 43  
Db 6103 IADGKKRVLVKDATLQDMGTYYVMVGAARAAHLTVIEKLRIVVPLKDRVKEQEVVF 6162  
QY 44 ----- 43  
Db 6163 NCEVNTGAKAKWFRNEBAIFDSSKYIILOKDLVYTLIRDAHLDDQANVNSLNRHGE 6222  
QY 44 ----- 43  
Db 6223 NVKSAANLIVEEDLRIVEPLKDIETMEKKSFTFWCKVYNRLNVLTKWTKNGEEVFPDNRV 6282  
QY 44 ----- 43  
Db 6283 SYRVDKYKHLTIKDCGFPDEGEYIVTAGQDKSVAELLIIEAPTEFVEHLEDQTVTEFDD 6342  
QY 44 ----- 43  
Db 6343 AVFSCQLSREKANWKYRNGREIKEGKKYKFEKOGSIHRLIIKDCRLDDECEYACGVDR 6402  
QY 44 ----- 43  
Db 6403 KSRARLFVEEIPVEIIRPPQDILEAPGADVFLAELNKDKVEVQWLRNNWVVQGDKHQM 6462  
QY 44 ----- 43  
Db 6463 MSEGKIHLRIQIDIKPRDQGEYRFTAKDKARAKLELAAAPKIKTADQDLVVDVGKPLTM 6522  
QY 44 ----- 43  
Db 6523 VVPYDAYPKAEAEWFKENEPLSTKTIDTAQTSFRILEAKKGDKGRYKIVLQNHKGAE 6582  
QY 44 ----- 43  
Db 6583 GFINKLVIDVPGVRNLEVTFETFDGEVSLAWEEPLTDGSKIIIGYVVERDRIKRTWVLA 6642  
QY 44 ----- 43  
Db 6643 TDRASCEFTVTGLQGGVEYLFVRSAARNVGTGEPVETDNFVEARSKYDVPGLNVTI 6702  
QY 44 ----- 43  
Db 6703 TDVNRFGVSLTWPEPYDGGAEITNYVIELRDKTSIRWDMTAMTVRAEDLSATVTDVVEGQ 6762  
QY 44 ----- 43  
Db 6763 EYSFRVRAQNRIGVGKPSAATPFVKVADPIERPSPPNLTSSDQTOSSVOLKWEPLKDG 6822  
QY 44 ----- 43

Db 6823 GSPILGYIIERCEGKNWIRCNMKLVLPETYKVTGLEKGNKYLYRVAENKAGVSDPSE 6882  
QY 44 ----- 43  
Db 6883 ILGPLTADDAFVEPTMDLSAFKDGLEIVVPNPITILVPSTGYPRPTATWCFGDKVLETCG 6942  
QY 44 ----- 43  
Db 6943 RVKMKTLNAYAEVLSPERSDKGIYTLKLENVKTISGEIDVNVNVIARPSAPKELKFGDI 7002  
QY 44 ----- 43  
Db 7003 TKDSVHLTWEPDDDDGSPLTGYVVEKREVSRTWTVMDFVTDLEFVPLDVOGKEYLF 7062  
QY 44 ----- PP----- 45  
Db 7063 KVCARNKCGPEPAYVDEPVNMSTPATVPDPENVKWRDRTANSIFLTWDPKNDGGSRI 7122  
QY 46 ----- CP----- 47  
Db 7123 KGYIVERCPRGSDKWVACGEPVAETKMEVTGLEEGKWYAYRVKTLNRQGAKSPRTEI 7182  
QY 48 ----- APELL----- 52  
Db 7183 QAVDTQEAPEIFLDVKLLAGLTVKAGTKIELPATVTGKPEKITWTKADMILKQDKRITI 7242  
QY 53 ----- GGPSVF----- 58  
Db 7243 ENVPKSTVTIVDSKRSDTGTIIIEAVNVCGRATAVVEVNVLDKPGPPAADIITDVTNES 7302  
QY 59 ----- 58  
Db 7303 CLLTWNPRDDGSKITNVVVERRATDSVWHKLSSTVKDTNFKATKLIPKEYIFRVAA 7362  
QY 59 ----- 58  
Db 7363 ENMYGAGEPVQASITAKYQFDPGPPTRLEPSDITKDAVTLTWCEPDDGGSPITGYW 7422  
QY 59 ----- 58  
Db 7423 ERLDPTDKWRCNMPVKDITYRVKGLTNKKYRFRVLAENLAGPGRFSKSTEPILLIKD 7482  
QY 59 ----- 58  
Db 7483 PIDPPWPGKPTVKDVKTSVRLNWTKPBHDGAKIESVYIEMLTGTDEWVRAEVSPT 7542  
QY 59 ----- LFPPKP----- 64  
Db 7543 TQHLPLGLMEGOEYSFRVRAVNKAGESESPSDPVLCREKLYPPSPRWLEVINITKNT 7602  
QY 65 ----- KDT----- LMSRT-- 73  
Db 7603 ADLKWTVPKGGSPITNIVIVEKRDVRRKGWOTVDTTVKDTCTCTVTPLTGEGSLYFVRVAA 7662  
QY 74 ----- PEYTCVVVDV----- 83  
Db 7663 ENAIGQSDYTEIEDSVLAKDTFTTGPYPVALAVDVTKRHRVLDKWEPPKNDGGRPIQRV 7722  
QY 84 ----- SHE 86  
Db 7723 IEKKERLGRWYKAGTAGPCDNCFRVTDVIEGTEVQFQVRAENEAGVGHPSPTIELSIE 7782  
QY 87 DB----- 88  
Db 7783 DPTSPSPPLDLHVTDAGRKKHTAIAWKPPKNGGSPIIIGYHVEMCPVGTEKMWVRNRSRI 7842  
QY 89 -EVKF----- 92  
Db 7843 KDLKFVEEGVVPDKYEVLRVRAVNAIGYSEPSISENVVAKDPCKPTIDLETHDIIVI 7902  
QY 93 ----- NWYVDG----- 98  
                  : : : : :  
                  : : : : :

Db 7903 EGEKLSIPVPPRAVPVPTVSHHKDGKEVKASDRLTMKNDHISAHLEVPKSVRADAGIYTI 7962  
Qy 99 ----- 98  
Db 7963 TLENKLSATASINVKVIGLPCPKDIKASDITKSSCLTWEPPFDGTPILHVLER 8022  
Qy 99 ----- 98  
Db 8023 EAGRRTYIPVMSGENKLSWTVKDLIPNGEYFFRVKAVNKVGGEYIELKNPVIAQPKOP 8082  
Qy 99 ----- 104  
      VEVHNA-----  
      |||||  
Db 8083 PDPPVDVEVHNPTAEAMTITWKPPPLYDGGSKIMGYIEKIAKEERWKRCHLVPILTY 8142  
Qy 105 ----- 104  
Db 8143 TAKGLEEKEXQFVRAENAAGISEPSRATPTTKAVDPIDAPKVILRTSLEVKRGDEIAL 8202  
Qy 105 ----- 104  
Db 8203 DASISGSPYTTITWIKDENIVPBEIKKRAAPLVRRRKGEYOEPEFVPLTLQRLSIDNS 8262  
Qy 105 ----- 104  
Db 8263 KKGESQLVRDSLDPDHGLYMIKVENDHGIAKAPCTVSLVDTPGPPINFVPFEDIRKTSVL 8322  
Qy 105 ----- 108  
      -----KTKP-----  
      |||||  
Db 8323 CKWEPPLDGSGEINIYTLKKDKTKPDSEWIVVTSLRHCKYSVTKLIEGKEYLFRVRA 8382  
Qy 109 ----- 108  
Db 8383 ENRFGPGPCVSKPLVAKDPFGPDADPKPIVEDVTSNMLVKWNEPKDNGSPILGYWLE 8442  
Qy 109 ----- 108  
Db 8443 KREVNSTHWSRVNKSLLNALKANVDLGLGLTYFVRCAENAAAGPGRFPPSPDKTAHP 8502  
Qy 109 ----- 108  
Db 8503 ISPPGPIPRVTDTSITIELEWEPFNGGEIGVGVFDQLVGTNKNWSRCTEKMIKYR 8562  
Qy 109 ----- 108  
Db 8563 QYTVKEIREGADYKLRVSAVNAAGEPPGETQPVTVAEPPQPPAVELDVSVKGGIIMAG 8622  
Qy 109 ----- 108  
Db 8623 KTLRIPAVVTGRPVTKVWTKEGELOKDRVVIDNVGTKSELIITKDALKRDKHGRYVITAT 8682  
Qy 109 ----- 108  
Db 8683 NSCGSKFAAARVEVDVPGVPLDLKPVVTRNKMCLLNSDPEDDGGSEITGFIIERKDAK 8742  
Qy 109 ----- 108  
Db 8743 MHTWRQPIETERSKCDITGLGEOYKFRVIAKNKFGCPVEIGPILAVDPLGPPTSPE 8802  
Qy 109 ----- 108  
Db 8803 RLTYTERQRSTITLDWKEPRNSGSPIOGYIIIEKRRHDKPDFRVNKRCLPTTSFLVENL 8862  
Qy 109 ----- 108  
Db 8863 DEHQMYEFRKAVNAGESEPSPLNVVIOQDEVPPTIKLRLSVRGDTIKVKAGEPVHIP 8922  
Qy 109 ----- 108  
Db 8923 ADVTGLPMPKIEWSKNETVIEKPTDALQITKEYSRSEAKTELSIPKAVREDKGTYTVA 8982  
Qy 109 ----- 108  
Db 8983 SNRLGSVFRNVHVEYDRPSPRNLAVIDIKAESCYLTWDAPLDNGSGSEITHYVIDKRDA 9042

Qy 109 ----- 108  
Db 9043 SRKKAWEVINTAVEKRYGIWKLIIPNGOYEFRRVAVNKYGISDBECKSDKVVIQDPYRLP 9102  
Qy 109 ----- 108  
Db 9103 GPPGPKVLARTKGSMLYSWTPLDNGGSPITGYWLEKREEGSPYWSRVSRAPITKVGLK 9162  
Qy 109 ----- 108  
Db 9163 GVEFNVPRLLEGVYQFRAMAINAAGIGPPSEPSDEVAGDPIFPPGPPSCEVSKDKTKS 9222  
Qy 109 ----- 108  
Db 9223 SISLGWKPAPKGGSPIKGYIIVEMQEEGTTDMKRVNEPDKLITTCCEVVPNLKELRKYRF 9282  
Qy 109 ----- 108  
Db 9283 RVKAVNEAGESEPSDTTGEIPATDIOEBEVEFIDIGAQDCLVCKAGSQIRIPAVIKGRPT 9342  
Qy 109 ----- 108  
Db 9343 PKSWEFDGKAKAMKDGVDIPEDAQLETAENSVIIIPECKRSHTKYKSITAKNKAGQ 9402  
Qy 109 ----- 108  
Db 9403 KTANCRVKVMDVPGPDKLVSDITRGSCRLSWKMPDDDDGDRINKGYVIEKRTIDGKAWT 9462  
Qy 109 ----- 113  
      -----REEQY-----  
      |:|:|  
Db 9463 KVNPDGCGSTTFVWPDLLSEQOYFFRVAENRFGIGPPVETIORTTARDPIYPPDPPIKLK 9522  
Qy 114 ----- 113  
Db 9523 IGLITKNTVHLSWKPPKNDGGSPVTHYVECLAWDPTGTKEAWRQCNKRDEVELOFTVE 9582  
Qy 114 ----- 113  
Db 9583 DLVEGGEYFRKAVNAAGVSKPSATVGPCDCORPDMPPSIDLKEFNEVEEGTWNVIVAK 9642  
Qy 114 ----- 113  
Db 9643 IKGVFPPTLTWFKAPKKPNKPNKPELVYDTHVNKLWVDDTCTLVIPQSRSDTGLYTITAV 9702  
Qy 114 ----- 113  
Db 9703 NNLGTASKEMRLNVLGRPPVGPVPIKFESVSADQMTLSWFFPKDDGGSKITNYVIEREA 9762  
Qy 114 ----- 113  
Db 9763 NRKTWVHVSSEPKECTYTIKLLSGHEVVFPRMAQNKYIGICEPLDSEPETARNLFSVPGA 9822  
Qy 114 ----- 113  
Db 9823 PDKPTVSSVTRNSMTVNWEEPEYDGGSPVTGYWLEMKDITTSKRWNKRVNRDPKAMTLGVS 9882  
Qy 114 ----- 113  
Db 9883 YKVTGLIEGSDYQPRVYAINAAGVPASLPSPDATARDPIAPGPPPKVTDWTKSSADL 9942  
Qy 114 ----- 113  
Db 9943 EWSPPKLDGGSKVTGYIIVEYKEBKEBKEKKEVRGRTKLVTGLKEGAFYKFRVSAVN 10002  
Qy 114 ----- 113  
Db 10003 IAGIGEPVTDVTEMKDRLVSPDLQDASVRDRIVVHAGVIRIIAYVSGKPPPTWTN 10062  
Qy 114 ----- 113  
Db 10063 MNERTLQPEATIETAISSSMWIKNCORSQHGVSLLAKNAGEKKTIIIVDLDVDPGV 10122

QY 114 ----- 113  
Db 10123 GTPFLAHLNLTNESCCLTWFSPEDDGGSPITNVYIEKRESRRRAWTPVTVTVTRQNAVQ 10182  
QY 114 ----- 113  
Db 10183 LIQKAYFFRIAAENSIGMPFVETSEALVIREPITVPERPELVEKVTNKTVTLTWP 10242  
QY 114 ----- 113  
Db 10243 PKYDGSSEIINYLESRLIGTEKFHKVTNDNLLSRKYTVKLGKEDTYEYRVSANNVQ 10302  
QY 114 ----- 113  
Db 10303 GRPSCFKPITCKDELAPPTLHLDFRDKLTIRVGEAFALTGRYSGKPKVSWFKDEADV 10362  
QY 114 -----NST----- 116  
|||  
Db 10363 LEDDRTHIKTTPATLALEKIKAKRSDSGKYCVVVENSTGSRGFCOVNVVDHPGPPGV 10422  
QY 117 ----- 116  
Db 10423 SFDEVTKDYMWISWKPLDDGGSKITNYIIEKKEGVDMVMPVTSASAKTTCVSKLLEG 10482  
QY 117 ----- 116  
Db 10483 KDYIFRIHAENLYGIDPLVSDSMKAKDRFPDAPQPIVTEVTKDSALVTWKNPHDGG 10542  
QY 117 ----- 116  
Db 10543 KPITNYILEKRETMKRWARVTKDPIHPYTKFRVPDLLEGQYFRVSAENEIGIDPSP 10602  
QY 117 ----- 116  
Db 10603 PSKPVFAKPIAKSPVNPALDITCNSVOLTWQPRHDGSKILGIYVEYQKVQDEW 10662  
QY 117 -----YRVSV----- 122  
:|:|:  
Db 10663 RRANHTPESCETKYKVTGLRDGQYKFRVLAVNAAGESDPAHVPEPVLVKDRLEPELI 10722  
QY 123 ----- 122  
Db 10723 LDANWAREOHKIVGDTLRLSAILIKGVFPKVTWKEDRDAPTAKRIDVTPVGSKLEIRNA 10782  
QY 123 -----LTV----- 125  
|||  
Db 10783 AHEDGGIYSLTVENPAGSKTVSKVLVLDKPGPPRDLEVSEIRKDSCLYLTWKEPLDDGS 10842  
QY 126 ----- 125  
Db 10843 VITNYVVERRDVASQWSPLSATSKKSHFAKHLNEGNOYLFRAVAENOYGRGPFVETPK 10902  
QY 126 -----LH-----ODWL-- 131  
|||  
Db 10903 PIKALDPLHPGPPKDLHVDVDKTEVSLVWNKPDGGSPTIGYLVEYQEBGTQDWIKF 10962  
QY 132 ----- 131  
Db 10963 KWTNLECVVTGLQOGKTYFRVRAENIVGLGLPDTTPIECQEKLVPPSVELDVKLI 11022  
QY 132 ----- 131  
Db 11023 LVVAGKTVRFPAILIRGVPTAKWTTDGEIKTDEHYTVETDNFSSVLTIKNCLLRDTG 11082  
QY 132 ----- 131  
Db 11083 EYQIIVSNAAGSKTVAVHLTVLDVPGPTGPINILDVTPHEMTISWQPKDGGSPVINY 11142  
QY 132 ----- 131  
Db 11143 IVEKODTRKDTWGVSSGSSKTKLIPHLOKCEYFRVRAENKIGVGPPLDSTPTVAKH 11202  
QY 132 ----- 131

Db 11203 KESPPSPCKPVVTDITENAAVSWTLPKSDGGSPITGYMERREVTGKVRVKNKPIAD 11262  
QY 132 ----- 131  
Db 11263 LKFRVTGLYEGNTYFRVFAENLAGLSKPSSPSDPIKACRPKPKPPGPPINPKLKDSRET 11322  
QY 132 ----- 131  
Db 11323 ADLVWTKPLSDGGSPILGYVVECOQKPTAQWNRINKDELIRQCAFRVPGPLIEGNEYFR 11382  
QY 132 ----- 131  
Db 11383 KAANIVGEGEPRELAESVIAKDLHPPEVELDVTCDVITVRVGQTIIRILARVKGPEPD 11442  
QY 132 ----- 131  
Db 11443 ITWTKEGKVLVREKRVLDLQDLPRVELOKEAVRADHGKYYIISAKNSSGHAQSAIVNVL 11502  
QY 132 ----- 131  
Db 11503 DRPGPCQNLKVTNVTKENCTISWENPLDNGGSEITNFIVEYRKPNOGWSIVASDVTKRL 11562  
QY 132 ----- 131  
Db 11563 IKANLLANNEYFRVCAENKVGVPPTIETKTPILAINPIDRPGEPENLHIADKGTFFVL 11622  
QY 132 ----- 131  
Db 11623 KWRRPYDGGSPNLSYHVERRLKGSDDERVHKGSIKETHYMWVDRVCENOIYEFVRVOTKN 11682  
QY 132 ----- 131  
Db 11683 EGGESDWKTEEVVVKEDLQKPVLDLKLSQLTVKAGDTIRLEAGYRGKPFPEVAWTKDK 11742  
QY 132 ----- 131  
Db 11743 DATDLTRSPRVKIDTRADSSKFSLTAKRSDGGKYVVTATNTAGSEFAYATVNVLDKPGP 11802  
QY 132 ----- 131  
Db 11803 VRNLKIVDVSSDRCTVCWDPPEDDGGCEIQNYILECETKRMVWSTYSATVLTPTGTVTR 11862  
QY 132 ----- 131  
Db 11863 LIEGNEYIFRVAENKIGTGPPTESKPVIAKTKYDKGPRDPPEVTKVSKEEMTVWNPP 11922  
QY 132 ----- 131  
Db 11923 EYDGGKSITGYFLEKKEKHSTRWVPVNVKSAIPERRMKVONLLPDHEYQFRVKAENEIG 11982  
QY 132 ----- 131  
Db 11983 EPSLSPRPVAKDPIEPPGPTNFRVVDTKHSITLGMGKPVYDGGAPIIGYVEMRPKI 12042  
QY 132 ----- 131  
Db 12043 ADASPDGKWKRCNAAQLVRKEFTVTSLDENOEYEFVCAQNOVGIGRPAELKEAIKPRE 12102  
QY 132 ----- 131  
Db 12103 ILEPPEIDLDSMRKLVIVRAGCPIRLFAIVRGPAPKVTWRKVGDVNVVRKGOVDLVD 12162  
QY 132 ----- 131  
Db 12163 MAFLVTPNSTRDDSGKYSLTLYNPAGEKAVFVNVRLDTPGPVSDLKVSVDVTKTSCHVSW 12222  
QY 132 ----- 131  
Db 12223 APPENDGGSQVTHYIVEKREADRKTWSTVTPVKKTSFHVTLNPGNEYFRVTAVNEVG 12282  
QY 132 ----- 131

Db 12283 PGVPTDVPKPVLASDPLSEDPDRKLEATEMTKNSATLAWPLPLRGGAKIDGYIISYRE 12342  
Qy 132 ----- 131  
Db 12343 EQPADRWTEYSVKWDLVLVVTGLKEGKKYFRVAARNAGVSLPREAGVYEAKEQLP 12402  
Qy 132 ----- 131  
Db 12403 PKILMEQITIKAGKLRIEAHVYCKPHTCKWKKGEDENVYSSHLAVHKADSSILIK 12462  
Qy 132 ----- 131  
Db 12463 DVTRKSGYSLTAENSSGTDTKIKVVMDAPGPPQPFDISDADACSLSHWPLED 12522  
Qy 132 ----- 131  
Db 12523 GGSNTNIVEKCDVSRGOWTALASVTKTSCRVGKLLPGQBYIFRVAERNFGISEPLT 12582  
Qy 132 ----- 131  
Db 12583 SPKMQAOPFGVPSEPKNARTKVNKDCIFVANDRPDSGGSPILIGYLLIERKERNSLMW 12642  
Qy 132 ----- 131  
Db 12643 KANDTLVRSTYPCAGLVEGLEYSFRIYALNKAGSSPPKTEYVTARMPVDPGKPEVI 12702  
Qy 132 ----- 131  
Db 12703 DVTKSTVLIWARKHGGSKLIIGFVEACKLPDGKWRWCNTAPHQIPOEYATATGLEEK 12762  
Qy 132 ----- 131  
Db 12763 AOYQRAIARTAVNISPPSPDPVTILAENVPPRIDLSVAMKSLITVKAGTNVCLDATV 12822  
Qy 132 ----- 131  
Db 12823 FGKPMPTVSWKKGDTLLKPAEGIKMAMORNLCLELFSVNRKDSGDTITTAENSSGSKSA 12882  
Qy 132 ----- 131  
Db 12883 TIKLVLDKPGPASVKINKMYSDBRAMLSWEPPLDGGSEITNYIVDKRETSRPNWAQVS 12942  
Qy 132 ----- 131  
Db 12943 ATVPITSGVEKLIBGHEYQFICAENKYGVDPVFTEPAIAKNPYDPGRCDDPPVISNI 13002  
Qy 132 ----- 131  
Db 13003 TKDHMTVSWKPPADGGSPITGYLLEKRETQAVNWKVNRKPIIERTLTKATGLQEGTEYE 13062  
Qy 132 ----- 131  
Db 13063 FRVTAINKAGPKPSDASKAAYARDPOYPAPPAPPKYVDITRSVSLSWGKPAYDGGSP 13122  
Qy 132 ----- 131  
Db 13123 IIGYLVEKRADSNWRCNLPNLQKTRFEVTGLMEDTQOFRVYAVNKIGYSDPSDVP 13182  
Qy 132 ----- 131  
Db 13183 DKHPKDIILIPPEGEHDADRLKTLILRAGVTMRLYVPVKGRPPKITTWSKPNVNLDRIG 13242  
Qy 132 ----- 131  
Db 13243 LDIKSTDFDIFLCENVNKYDAGKYILTLNENSCGKKEYIIVVKVLDTPGPPINVTVKEIS 13302  
Qy 132 ----- 131  
Db 13303 KDSAVVTWEPIIDGGSPIINYVQKRAERKSWSTVTECSKTSFRVPNLEEGSKYFFR 13362  
Qy 132 ----- 131  
Db 13363 VFAENEYIGDPGETRDAVKASQTFPGVPVVDLKVRSVSKSCSIGWKKPHSGGSRIGYV 13422

Qy 132 -----NGKEY----- 136  
Db 13423 VDFLTEENKQWQVMKSLSLQYSAKDLTEGKEYTFRVAENENGEPTSEITVWARDVVA 13482  
Qy 137 ----- 136  
Db 13483 PDLDLKGLPDLCLYLAKENSFRLKIPKGPAPSVSWKKGEDPLATDTRVSVESAVNTT 13542  
Qy 137 ----- 136  
Db 13543 LIVYDCQKSDAGKYITITLKNVAGTKEGTISIKWVGKPGIPTGPIKFEDEVTAEMTLKWP 13602  
Qy 137 ----- 136  
Db 13603 PKDDGSEITNYILEKRSVNKNWVTCASAVQKTTFRVTRLHEGMEYTFRVAENKYGVG 13662  
Qy 137 ----- 136  
Db 13663 EGLKSEPIVARHPEDVPDAPPPPNIVDVHDSVSLTWTDPKKTGGSPITGYHLEFKERNS 13722  
Qy 137 ----- 136  
Db 13723 LLWKRANKTPIMRDKFKVTGLTEGLEYEFRVMAINLAGVGPSPSEPVALDPIPPGK 13782  
Qy 137 ----- 136  
Db 13783 PEVINITRNSVTLIWTEPKYDGGHKLTCYIVEKRDLPSKSHWKANHVNVPECAFTVTDLV 13842  
Qy 137 ----- 136  
Db 13843 EGGKYEPRIRAKTAGAISAPSESTETIICKDEYEAPTVLDPTIKDGLTIKAGDTIVLN 13902  
Qy 137 ----- 136  
Db 13903 AISILGKPLPKSSWSKAGDIRPSDITQITSTPTSSMLTIKYATRKDAGEYITATNPF 13962  
Qy 137 ----- 136  
Db 13963 TKVEHVKTVDLVPDGPVPEISVNSABKATLTWTPPLEDDGGSPKSVILEKRETSRLW 14022  
Qy 137 ----- 136  
Db 14023 TVSEDIQSCRHVATKLIQNGNEYIFRVSANVHYKGEPVQSEPKVMVDFGPPGPKPE 14082  
Qy 137 -----KCKVSN----- 142  
Db 14083 VSNVTKNTATVSWKRPVDDGGSEITGYHVERREKKSRLWRRAIKTPVSDLRCKVTGLQEG 14142  
Qy 143 ----- 142  
Db 14143 STYEFVRSAENRAGIGPPSEASDSVLMKDAAYPPGPPSNPHVTDITTKKSASLANGKPHYD 14202  
Qy 143 ----- 142  
Db 14203 GGLEITGYVVEHQVKGDEAWIKDGTGTALRITQFVVPDLOTKEYNFRISAINDAGVGP 14262  
Qy 143 -----KALPAP----- 148  
Db 14263 AVIPDVEIVEREMAPDFELDAELRRTLVRAGLSIRIFVPIKGRPAPEVTTWKONINUKN 14322  
Qy 149 ----- 148  
Db 14323 RANIENTESFTLLIIPECNRYDTGKFVMTIENPAGKSGFVNVRLDTPGPVNLRLPTDI 14382  
Qy 149 -----IEK----- 151  
Db 14383 TKDSVTLHWDLPIDGGSRITNYIVEKREATRKSYSTATTCKHCKTYKVTGLSECEYFF 14442  
Qy 152 ----- 151  
Db 14443 RVMAENEYIGEPTEPTEPVPKASEAPSPDLSINIMDITKSTVSLAWPKPKHGGSKITGY 14502



QY 152 ----- 151  
Db 14503 VIEAQRKSDQWTHITVTKGLECVVRNLTEGEYTFQVMVANSAGRSAPRESPIVKEQ 14562  
QY 152 ----- 151  
Db 14563 TMLPELDLGGIYQKLVIAKAGONIKVEIPVLGRPKFTVTWKGDQILKOTQRVNFETTAT 14622  
QY 152 ----- 151  
Db 14623 STILNINECVSRSDGYPPLTARNIYGEVGDVITIQVHDIPGPGTGPDKFDEYSSDFVTF 14682  
QY 152 ----- 151  
Db 14683 WDPENDGGVPISNVYVEMROTDSTTWELATTIVIRTTYKATRLTTGLEYQFRVKAQRY 14742  
QY 152 ----- 151  
Db 14743 GVGPGITSAMIVANYPFKVPGPPTQVTAVTKDSMTISWHEPLSDGGSPILGYHYVERKE 14802  
QY 152 ----- 156  
                  TISKA-  
                  I:III  
Db 14803 RNCILMQTVSKALVPGNIFKSSGLTDGIAEFVRVIAENMAGKSKPSPMLALPIDP 14862  
QY 157 ----- 156  
Db 14863 PGKPVPLNITRHTVTLKWAKEPYTGFKITSYIVEKRDLPNGRWLKANFENILENEFTVS 14922  
QY 157 ----- 156  
Db 14923 GLTEDAAYEFVRVIAKNAAGISPPSEPSDAITCRDDVEAPKIKVOVKDVTILKAGEAF 14982  
QY 157 ----- 156  
Db 14983 RLEADYSGRPPPTMWSKDGKELEGTAKLEIKIADFTNLVKNKDSRRDGGAYTLTATNP 15042  
QY 157 ----- 156  
Db 15043 GGFAKHIFNVKVLDRPGPEGLAVTEVTSEKCVLSWFPPDDGAKIDHYIVOKRETSR 15102  
QY 157 ----- 156  
Db 15103 LAWTVASEVQVTKLVTKLLKAGNEYIFRMAVNKYGVGELESEPVLAENVYGPDPKP 15162  
QY 157 ----- 156  
Db 15163 NPEVTTITKDSMVVCHPDSDDGSELIINYIVERRDKAGQWIKCNKKTILDLRYKVSGL 15222  
QY 157 ----- 156  
Db 15223 TEGHEYEFRIMAENAGISAPSPSYKACDVFYKPGPNRVLDTSRSSISIAWKP 15282  
QY 157 ----- 156  
Db 15283 IYDGSSEITGYMVEIALPEDEWQIVTPAGLKATSYITIGLTENOYKIRIYAMNSEGL 15342  
QY 157 ----- 156  
Db 15343 GEPALVPGTPKAEDRMLPPEIELDADLRKVVTIRACCTLRFPVFKGRDPEVKWARDHG 15402  
QY 157 ----- 156  
Db 15403 ESLDKASIESASYTLIIVGNVNRFDGKYLTVENSNGSKSAFVNVRLDTPGPQDLK 15462  
QY 157 ----- 156  
Db 15463 VKEVTKTSVTLWDPPPLDGGSKIKNYIVEKRESTRKAYSTVATNCHKTSWKVDLQEGC 15522  
QY 157 ----- 156  
Db 15523 SYIFRVAENEYIGLPAETASVKASERPLPGKITLMDVTRNSVLSWEKPEHDGGR 15582  
QY 157 ----- 156

Db 15583 ILGYIVEMQTKGSKWATCATVKVTEATITGLIQGEEYFVRVSAQNEKGISDRQLSPV 15642  
QY 157 ----- 156  
Db 15643 IAKDLVIPPAFKLLFNFTFVLAGEDLKVDVPFICRPTPAVTWHDNDVPLKQTRVNAEST 15702  
QY 157 ----- 156  
Db 15703 ENNSLLTIKDACREDVGHVVKLTNSAGEAIELNVLIDKPGPTGPVKMDEVADTSIT 15762  
QY 157 ----- 156  
Db 15763 LSWGPKYDGGSSINNYIVEKRDSTTTWQIVSATVARTIKACRLKTCGEYQFRIAAEN 15822  
QY 157 ----- 156  
Db 15823 RYCKSTYLNSEPTVAQYPPKVPGPPTPVVTLSSRDSMEVQWNEPISDGGSRVIGHLER 15882  
QY 157 ----- 156  
Db 15883 KERNSILWVKLNKTIPIQTKFKTTGLEEGVEYEFVRVSAENIVGIGKPSKYSECYVARDPC 15942  
QY 157 ----- 156  
Db 15943 DPEGRPEALIVTRNSVTLQWKPTYDGGSKITGYIVEKKELPEGRWMAKSTNIIDTHFE 16002  
QY 157 ----- 156  
Db 16003 VTGLVEDHRYEFVRVIAARNAAGVSEPSSESTGAIATARDEVDPPIRISMDPKYKDIVVHAGE 16062  
QY 157 ----- 156  
Db 16063 SFKVDADIYKPIPTIOWIKGDOELSNATLEIKSTDFTATSLVKDAVRVDSNGYILKAK 16122  
QY 157 ----- 156  
Db 16123 NVAGERSVTNVKVLDRPGPPEGVVIVSGVTAECTLAWKPLQDGGSDIINIVERRET 16182  
QY 157 ----- 156  
Db 16183 SRLVTVVDANVOTLSCKVTKLLEGNEYTFRIMAVNKNYGVGELESEPVVAKNPFVVPDA 16242  
QY 157 ----- 156  
Db 16243 PKAPEVTITKDSMIVVWERPASDGGSEILGYVLEKXDEKGIWTRCHKRLIGELRLVT 16302  
QY 157 ----- 156  
Db 16303 GLIENHDYEFVRVSAENAGLSEPPSPAYQKACDPIYKPGPPNNPKVIDITRSSVFLWS 16362  
QY 157 ----- 156  
Db 16363 KPIYDGGCEIQGYIVEKCDVNVGEMTCTPTGINKNTNIEVEKLEKHEYNFRICAINKA 16422  
QY 157 ----- 156  
Db 16423 GVGEHADVPPIIVEKLEAPDIDLDELRLKIINIRAGSLRLFPVPIKGRPTPEVKWGV 16482  
QY 157 ----- 156  
Db 16483 DGEIRDAAIIDVTSSFTSLVDNVRVDSGKYTLTLENSSGTSKSAFVTVRVLDTPTSPPVN 16542  
QY 157 ----- 156  
Db 16543 LKVEITKDSVSTWEPPLDGGSKIKNYIVEKREATRKSAAVVTNCHNSMKIDQLQE 16602  
QY 157 ----- 156  
Db 16603 GCSYIFRVAENEYIGLPAQTAADPIKVAEVPQPPGKITVDDVTRNSVLSWTKPEHDGG 16662  
QY 157 ----- 156

Db 16663 SKIIQIIVEMQAKHSEKSEKARVKSLOAVITNLITQGEYLFRRVVAVNEKGRSDPRSLAV 16722  
QY 157 ----- 156  
Db 16723 PIVAKDLVIEPDVKPAFSSYSVOVGDLKMEVPISGRPKPTITWTKDGLPLKQTRINTV 16782  
QY 157 ----- 156  
Db 16783 DSLDLTLLSIKETHKDDGGQYGITVANVVGQKTASIEIWLTDKPPKGPVKFDDVSAES 16842  
QY 157 ----- 156  
Db 16843 ITLSNPPPLYTGGCQITNIVOKRDTTITVNDVVSATVARTLKVTKLTGTGYQFRIFA 16902  
QY 157 ----- 156  
Db 16903 ENRYGQSALESDDPIVAQYYPKEPGPGTFFATAISKDSWVQWHEPVNNGSGPVIGHL 16962  
QY 157 ----- 156  
Db 16963 ERKERNILWTKVNTIITHTQFAQNLEEGIEYFRVVAENIVGVGKASKNSECYVARD 17022  
QY 157 ----- 156  
Db 17023 PCDPPTPEPIMVKRNEITLQWTKPVYDGGSMITGYIVEKKRDLDPGRWMAKSFNTVETQ 17082  
QY 157 ----- 156  
Db 17083 FTVSGLTEDQRYEFRVIAKNAAGAIKSPDSTGTPITAKDEVELPRISMDPKFROTIVNA 17142  
QY 157 ----- 156  
Db 17143 GETFRLEADVHGKPLPTIEWLGRDKIEBESARCEIKNTDFKALLIVKDAIRIDGGQYILR 17202  
QY 157 ----- 156  
Db 17203 ASNVAGSKSPVNVKVLDRPPEPGPVQVTGVTSEKSLTWSPPLODGGSDISHYVVEKR 17262  
QY 157 ----- 156  
Db 17263 ETSRLAWTVADEVVNTNSLKVTKLEGNEYFRIMAVNKYGVGEPLASAPVLMKNPFVLP 17322  
QY 157 ----- 156  
Db 17323 GPPKSLEVTNTAKDSMTVCWNRPSDGGSEIIGYIVEKKRDRSGIRWIKCNKRRTDLRLR 17382  
QY 157 ----- 156  
Db 17383 VTGLTEDHEYFRVSAENAGVGEPSPATVYKACDPVKPGPPTNAHIIVDTTKNSITLA 17442  
QY 157 ----- 156  
Db 17443 WCKPIYDGGSEILGYVVEICKADEEWQIVTPQTGLRVTREISKLTEHQEKIRVCALN 17502  
QY 157 ----- 156  
Db 17503 KVLGEATSVPGTVKPEDKLEAPELDSELKRGIVRVAGGSARIHIPFKGRPMPEITWS 17562  
QY 157 ----- 156  
Db 17563 REEGFTDKVQIEKGVNTQLSIDNCDNRNDAGKYILKLENSGSKSAFVTVKVLDTGPP 17622  
QY 157 ----- 156  
Db 17623 QNLAVKEVKDSAFVLEWPEPIIDGKAKVKNYVIDKRESTRKAYANVSSKCSKTSFKVENL 17682  
QY 157 ----- 156  
Db 17683 TEGAIYFRVMAENBFGVGPVETVDVAKAAEPPSPGKVTLTDVQSOTSASLWKEKPEHD 17742  
QY 157 ----- 156  
Db 17743 GGSRLVGVVEMQPKGTEKWSIVAESKVCNAVVTGLSSGQYQFRVAKYNEKKGSDPRVL 17802

QY 157 ----- 156  
Db 17803 GVPVIAKDLTITQPSLKLFPNTIYSIOAGEDLKIEIPVIGRPRNISWVKDGEPLKQTRVN 17862  
QY 157 ----- 156  
Db 17863 VEETATSTVLHIKEGKDDFGKYTVTATNSAGTATENLSVIVLEKPGVPVGRFDEVSA 17922  
QY 157 ----- 156  
Db 17923 DFVVISWEPPAYTGGCQISNYIVEKRDTTTTTHMVSATVARTTIKTKLTGTGYQFRI 17982  
QY 157 ----- 156  
Db 17983 FAENRYGKSAPLDSKAVIVQYFPKEPGPGTFFVTISISKDQMLVQWHEPVNDGGTKIIGY 18042  
QY 157 ----- 156  
Db 18043 HLEQEKNSILWVKNLKTPIQDTKFKTTGLDEGLEVEFKYSAENIVGIGKPSKVECFVA 18102  
QY 157 ----- 156  
Db 18103 RDCDPPGRPEAIYITRNNTLKWKKPAYDGGSKITGYIVEKKDLDPGRWMAKSFNTVLE 18162  
QY 157 ----- 156  
Db 18163 TEFTVSGLVEDQRYEFRVIAARNAGNFSEPSDSSGAIITARDEIDAPNASLDPKYKDVIV 18222  
QY 157 ----- 156  
Db 18223 HAGETFVLEADIRKPIPDVVVWSKDGKELEETAARMEIKSTIQKTTLVVKDCIRTDGQY 18282  
QY 157 ----- 156  
Db 18283 ILKLSNVGGTSSIPITVYKVLDRPSPGEPGLKVTGVTAEKCYLANNPLODGGANISHYII 18342  
QY 157 ----- 156  
Db 18343 EKRETSRLSWTQVSTEVQALNVKTKLLPGNEYIFRVMVANKYGIGEPLESGPVTACNPY 18402  
QY 157 ----- 156  
Db 18403 KPPGPPSTPEYSAITKDSMVVTWARPVDGGEIEGYILEKRDKEGVRWTRCNKNTLTDL 18462  
QY 157 ----- 166  
Db 18463 RLRTVGLTEGHSYEFERVAENAGVGEPSVIFYRACDALYPPGPPSPNPKVTDTSRSSV 18522  
QY 167 ----- 166  
Db 18523 SLAMSKPIYDGGAPVKGYVVEVKEAADEWTTCTPPTGLOGKQFTVTKLKENTENFRIC 18582  
QY 167 ----- 169  
Db 18583 AINSEGVGEATLPGSVVAQERIEBPPEIELDALRKVVVLRASATLRLFTVTKRPEPEV 18642  
QY 170 ----- 169  
Db 18643 KWEAKGILTDRAQIEVTSSTMLVIDNVRFPDGRYNLTLENNSGSKTAFVNVRLDSP 18702  
QY 170 ----- 169  
Db 18703 SAPVNLATIREVKDSVTLSEWPEPLIDGKAKITNYIVEKRETTKRAYATITNNCTKTTFRI 18762  
QY 170 ----- 169  
Db 18763 ENLOEGCSYFRVLASNEYIGLPAETTEPVVKVSEPPPLPGRVTLVDVTRNTATIKWEKP 18822  
QY 170 ----- 169  
Db 18823 ESDGSKITGYVVEVEMQKGEKSWSTCQVKTLEATISGLTAGEYVFRVAAVNEKGRSDP 18882

QY	170	-----	169
Db	18883	ROLGVPIARDIEIKPSVELPPHTENVKAREQLKIDVPFKGRQATVNNRKDGQTLKEIT	18942
QY	170	-----	169
Db	18943	RNVSSKTVTSLSIKEASKEDVGYELCVNSAGSITVPITIIIVLDRFPGPPIRIDEV	19002
QY	170	-----	169
Db	19003	SCDSITISNPPYDGGCCQISNVIVEKKEITSTTHIVSQAVARTSIKIVRLTGTSEYQF	19062
QY	170	-----	169
Db	19063	RVCAENRYKSSYESSAVVAEYFPFPGPGPKVVHATKSTMLVTWQVPVNDGGSRI	19122
QY	170	-----	169
Db	19123	GYHLEYKERSILWSKANKILIAADTOVKVSGLDEGLMEYRYVAENIAGIGKSKCEPV	19182
QY	170	PSRD-----	173
Db	19183	PARDCDPGQPEVTNITRKSVLSKSKPHYDGGAKITGYIVERRELDPGRWLKCNVTNI	19242
QY	174	-----	173
Db	19243	OETYFEVTELTEDQRYEYFRVAFARNADSVSESESTGPIIVKDDVEPRVMDVKFRDVI	19302
QY	174	-----	173
Db	19303	VKAGEVLKINADIAGRPLPVISWAKDGEIERARTEIISTDNHTLLTVKDCIRRDGTQ	19362
QY	174	-----	173
Db	19363	YVLTAKNAGTRSAVAVNCKVLDKPGPPAGPLEINGLTAEKCSLSWGRPBQDGGADIDYH	19422
QY	174	-----	173
Db	19423	RKKRETSHLAWTICEGELQWTSCKVTKLLKGNEYIFRVTGVNKYGVGEPLSAIKALDP	19482
QY	174	-----	173
Db	19483	FTVPSPPTSLEITSVTKESMTLWSPESDGGSEISGYIIERREKNSLRVVRNKKPVYD	19542
QY	174	-----	173
Db	19543	LRVKSTGLREGCEYEVRYAENAGLSLPSETSPLIARAEDPVFLPSPPKIVDSGKTT	19602
QY	174	-----	173
Db	19603	ITIAWVKPLFDGGAPITGYTVVEYKKSDDTDWKTISIQLRGTEYTIISGLTGAEYVFRKS	19662
QY	174	-----	173
Db	19663	VNVKASDPSDSDPOIAKEREEREEPLFDIDSEMRKTLIVKAGASFMTVPFRGRPVNVL	19722
QY	174	-----	173
Db	19723	WSKPDTLRTRAYVDTTDSRTSLTIENANENDSGKYTLTIQNVLSAASLTLVVKVLDTPG	19782
QY	174	-----	173
Db	19783	PPTNITVQDVTKESAVLSMDVPENDGGAPVKNYHIEKREASKKAWSVTNNCNRLSYKVT	19842
QY	174	-----	173
Db	19843	NLOEGAIYYFRVSGENEFVGIPAETKEGVKITEKPSPEKLGVTISISKDSVSLTWLKE	19902
QY	174	-----	173
Db	19903	HDGGSRLVHVVALEKGQKNWYKCAVAKSTHHVSVGLSENSEYFFRVAENOAGLSDPR	19962
QY	174	-----	173

Db	19963	ELLPLVLIKOLEPPEIDMKNFPSHTVYVRAGSNLKVDIPISGKPLPKVTLSDRGVPLKA	20022
QY	174	-----	173
Db	20023	TMRFNTEITAENLTINLKESVTADAGRYEITAANSSTTTAFINIVVLDLRGPPPTGPVI	20082
QY	174	-----	173
Db	20083	SDITESVTLKWEPKPYDGGSOVTNYILLKRETSTAVWTEVSATVARTMMKVMKLTGTGEE	20142
QY	174	-----	173
Db	20143	YQFRIKAENRFGISDHIDSACVTVKLPYTPPGPSTPWNTVNTRESITVGWHEPVSNGGS	20202
QY	174	-----	173
Db	20203	AVVGYHLEMKDRNSILWQANKLVIRTHFKVTTISAGLIYEFVRVYAENAGVKPSHPS	20262
QY	174	-----ELTKNOVSL-----	182
Db	20263	EPVLAIDACEPPRNRITDISKNSVLSWQOPAFDGGSKITGYIVERRDLPDGRWTKASF	20322
QY	183	-----TCL-----	185
Db	20323	TNVTETQFTISGLTQNSQYEFVRVAFARNAGVSGISNPSEVWGPITCIDSYGGPVIDPLEYT	20382
QY	186	-----	185
Db	20383	EVVKYRAGTSVKLRAGISGKPAPTIWYKDDKELQTNALVCVENTTDLASILIKDADRLN	20442
QY	186	-----	185
Db	20443	SGCYELKLRNAMASASATIRVQILDKPPGPGPIEFKTVTAEKITLLWRPPADGGAKIT	20502
QY	186	-----	185
Db	20503	HYIVEKRETSRVVWSVMVSEHLEECIITTKIIGNEYIFRVRVAVNKYGIGEPLESDSWA	20562
QY	186	-----	185
Db	20563	KNAPVTPGPGPIPEVTKITKNSMTVVWSRPIADGGSDISGYFLEKRDKSLGLWPKVLKET	20622
QY	186	-----	185
Db	20623	IRDRQKVTLGTENSQYRVCVAVNAAGQFPSEPSEYKAADPIDPPGPAKIRIADST	20682
QY	186	-----	185
Db	20683	KSSITLWSKPVYDGGSAVTGYVVEIRQGBEEWTTVSTKGEVTEYVYVSNLKPGVNYI	20742
QY	186	-----	185
Db	20743	FRVSAVNACQGEPIEMNEPVQAKDILEAPEIDLVAALRTSVIAKAGEDVQVLIPFKGRP	20802
QY	186	-----	185
Db	20803	PPTVTRKDEKLGSDARYSIENTDSSLLTIPOVTRNDTGKYLLTIENGVGEPKSTVS	20862
QY	186	-----	185
Db	20863	VKVLDTAACQKLVKHVSRGTVTLLWDPPLIDGGSPIINVIEKRDATKRTWSVSHKC	20922
QY	186	-----	185
Db	20923	SSTSFKLIDLSEKTPFFRVLAEINEIGIPEFCEETPEPVKAAEVPAPIRDLMSKDSTKTSV	20982
QY	186	-----	185
Db	20983	ILSWTRPDDGGSVITEYVVVERKKGQBTWSHAGISKTCIEVYSOLAEQSVLEPRVFAKN	21042
QY	186	-----	185

Db 21043 EKGLSDPVTIGTITVKELIITPEVDLSDIPGAQVTVRIGHNVHLELPYKGPSPISWLK 21102  
Qy 186 ----- 185  
Db 21103 DGLPLKESEFVRSKTEKNITILSIKNAKEHGGKYTVILDNAVCRIAIVITILGPPSK 21162  
Qy 186 ----- 185  
Db 21163 PKGPIRFDEIKADSVILSWDPEDNGGGEITCYSIEKRETSQTNKMKVCSVARTFKVP 21222  
Qy 186 ----- 185  
Db 21223 NLVKDAEQFRVRAENRYGVSOPLVSSIIVAKHOFRIIPGPGKPVIIYNTSDGMSLTWDA 21282  
Qy 186 -----VKGFY----- 190  
| | |  
Db 21283 PVDGSGSEVTFHVEKKERNLQKQVNTSPISGREYRATGLVEGLDYQFRVYAENSAGL 21342  
Qy 191 ---PSD----- 193  
| | |  
Db 21343 SSPSPDSKFTLAVSPDPGTPDYIDVTRETITLKNPPLRDLGGSKIYGYSIEKROGNER 21402  
Qy 194 ----- 193  
Db 21403 WVRNFTDVSECOYTVTGLSPGDRYEFRIIARNAVGTISPPSQSGIIMTRDENVPPIVE 21462  
Qy 194 ----- 193  
Db 21463 FGPEYFDGLIIKSGBSRLIKALVQGRPVPRVTWFKDGVGEIEKRMNMEITNVLGSTSLFVR 21522  
Qy 194 ----- 193  
Db 21523 DATDRHGVYTVAKNAGSASAKAEIKVKVQDTPGKVGPIRFTNITGKMTLWMDAPLND 21582  
Qy 194 ----- 193  
Db 21583 GCAPITHYIEKRETSRLAWALIEDKCAQYTAIKLINGNEYQFRVSAVKNFGVGRPLD 21642  
Qy 194 ----- 193  
Db 21643 SDPVVAQIQYTVDPAPGIPESNITGNSITLTWARPESDGGSEIQOYLIERREKKSTRWV 21702  
Qy 194 ----- 193  
Db 21703 KVISKRPISETRFRKVTGLTEGNEYEFHVMAENAGVSPGISRLIKREPVPNPGPTV 21762  
Qy 194 -----TAVW----- 198  
: : :  
Db 21763 .VKVTDTSKTTVLSLEWSKPVFDGMEIIGYIEMCKTDLDGMHKVNAEACVKTXYTVDLQ 21822  
Qy 199 ----- 198  
Db 21823 AGEYKFRVSAINGAGKDSCEVTGTIKAVDRLTAPELDIDANFKQTHVVRAGASIRLFI 21882  
Qy 199 -----ESN----- 201  
| | |  
Db 21883 AYQGRPTAVMSKPDNSLRLADIHTDTSFTLTVCNCRNDAGKYTLTVENNNGSKSI 21942  
Qy 202 ----- 201  
Db 21943 TFTVKVLDTPGPGPITPKDVTGRSATLMDAPLLDGGARITHYVVEKREASRRSQVIS 22002  
Qy 202 ----- 201  
Db 22003 EKCTRQIFKVNDAEGVPYFVRSVAVNEYGYEPEYEMPEPIVATEQPAPPRLDVVDTSK 22062  
Qy 202 ----- 201  
Db 22063 SSAVLAWLKPDDGSRITGYLLEMRQKSDLWVEAGHTKQLTFTVERLVEKTEYEFVRK 22122  
Qy 202 ----- 201  
Db 22123 AKNDAGYSEPREAFSSVIIKEPQIEPTADLTGITNOLITCKAGSPFTIDVPIGRPAKV 22182

Qy 202 ----- 201  
Db 22183 TWKLEEMRLKETDRVSITTTKDRUTTLTKVDSMRGDSGRYFLTLENTAGVKTFSTVWVIG 22242  
Qy 202 -----QOP----- 204  
| | |  
Db 22243 RGPVTPGPIEVSSVSAESCVLWSGEPKDDGGTEITNYIVKRESGTTAQLVNSSVKRTQ 22302  
Qy 205 ----- 204  
Db 22303 IKVTHLTKYMEYSFRVSSSENFVSKPLESAPIIAEHFPVPSPAPTRPREVHVHSANMSI 22362  
Qy 205 -----ENNYKTT----- 211  
| | |  
Db 22363 RWEOPYHDGSKIIGYWVEKKERNITLWVKENKVPCLCNKYKVTGLVEGLEYPFTYALN 22422  
Qy 212 ----- 211  
Db 22423 AAGVSKASEASRPIMAQNPDAPGRPEVTDVTRSTVSLWSAPAYDGGSKVVGYYIERKP 22482  
Qy 212 ----- 211  
Db 22483 VSEYDGRWLKCNVTIIVSDNFTVTALSEGDTYEFRLAKNAAGVISKSESTGPTVCRD 22542  
Qy 212 ---pp----- 213  
| |  
Db 22543 EYAPKAEELDARLHGLDVTIRAGSDLVLDAAVGGKPEPKIITWTKGDKELDLCEKVSLOYT 22602  
Qy 214 -----VLDSDG----- 219  
| | |  
Db 22603 GKRAVAVIKFCDRSDSGKYTLTVKNASGTKAVSMVKVLDSPGCGKLTVSRVTOEKCTL 22662  
Qy 220 ----- 219  
Db 22663 AWSLPQEDGGAETHYIVERRETSRLNWVIVEGECPTLSYVVYVRLKKNIEYFRVAVNK 22722  
Qy 220 -----SFFL----- 223  
| | |  
Db 22723 YGPGVPESEPIVARNSTIPSPCPIPEEVGTGKEHIIQWTKPESDGNESINLVOKR 22782  
Qy 224 ----- 223  
Db 22783 EKESLRWTRVKNKYVYDTRLKVTSMEGCDYQFRVTAVNAAGNSEPSESNFISCREPS 22842  
Qy 224 ----- 223  
Db 22843 YTPGPPSAPRVVDTTKHSISLAWTKPMYDGGTDIVGYVLEMQEKDQDQYRVHTNATRN 22902  
Qy 224 ----- 223  
Db 22903 TEFTVPDLKMGQYSFRVAANVKGMSYSESIAEIEPVERIEIPDLELADDLKKTVTIR 22962  
Qy 224 -----YSKLTVDK----- 231  
| | |  
Db 22963 AGASLRMLVSYGRPPPVITWSKOGIDLASRAIIDTTESYLLIIVDKVKNRYDAGKYTIEA 23022  
Qy 232 ----- 231  
Db 23023 ENQSGKKSATVLVKVYDTPGPCPSVKVEVSRDSVTITWEIPTDGGAPINNYIVEKREA 23082  
Qy 232 ----- 231  
Db 23083 AMRAFKTVTTKCKTLYRISGLVEGTMHYFRVLPENIYIGPCETSDAIVLSEVPLVPA 23142  
Qy 232 ----- 231  
Db 23143 KLEVVDVTKSTVTLAWEKPLYDGGSRLTGYVLEACACTERMMKVTLKPTVLEHTVSL 23202  
Qy 232 ----- 231  
Db 23203 NEGOYLFRIRAQNEKGVSEPRETVTAVTVQDLRVLPTIDLSTMPQKTIHVPAGRPVELV 23262

QY 232 ----- 231  
Db 23263 IPIAGRPAAASWFFAGSKLRESERVTHIKVAKLTIRETTIRDTGEYTLLELKNVTGT 23322  
QY 232 ----- 231  
Db 23323 TSETIKVIIIDKPPPTGIKIDEIDATSITISWEPPELGGAPLSGYVVEQDAHRPGW 23382  
QY 232 ----- 231  
Db 23383 LPVSESVTRSTFKFTRLTEGNEVFRVAATNRFPGISYLOSEVIECRSIRIPGPPETLQ 23442  
QY 232 ----- 231  
Db 23443 IFDVSRDGWTLTWYPPEDDGSQVTGYIVERKEVRADRVVRKNKVPVMTYRSTGLTSG 23502  
QY 232 ----- 231  
Db 23503 LEYEHRTAINARGSKPSRSPKPIVAMPDPIAPGKPNPRVTDTRTSVSLAWSPEDE 23562  
QY 232 ----- 231  
Db 23563 GSKVTGYLIEQKVDQHEWTKCNTPTTKIREYTLHLPOGAERYRVLACNAGGPGEPA 23622  
QY 232 ----- SRWQOQ- 237  
Db 23623 EVPGTVKVTEMLEYDPYELDERYQEGIFVRQGVIRLTPIKPKPPICKWTKEGODISK 23682  
QY 238 ----- 237  
Db 23683 RAMIATSETHELVIKEADRGDSGYDLVLENKCKGKAVYIKVRVIGSPNSPEPLEYDD 23742  
QY 238 ----- 237  
Db 23743 IQVRVRSWRPPADGGADILGYILERREVPKAAWTTIDSRVGTSLVVGLENVEYH 23802  
QY 238 ----- 237  
Db 23803 FRVSAENQFGISKPLKSEBPVTKPLNPPPEPPNPPELDVTKSSVLSWSRKPDDGGS 23862  
QY 238 ----- 237  
Db 23863 RVTGYIERKETSTDKVVRHNKTQITTTWYTVGLVPDAEYQFRIIAQNDVGLSETSPAS 23922  
QY 238 ----- 237  
Db 23923 EPVVKDPFDKPSQPGELEILSISKDSVTLQWEKPECDGKELGYWVEYRQSGDSAWKK 23982  
QY 238 ----- 237  
Db 23983 SNKERIKDKQFTIGGLLEATEYEFVFAENETGLSRPRRTAMSIKTKLTSGEAPGIRKEM 24042  
QY 238 ----- 237  
Db 24043 KDVTTKLGEAALSCQIVGRPLDIKWYRFGKELIQSRKYKMSDGRHTLTAVMTEEQED 24102  
QY 238 ----- 237  
Db 24103 EGYVTIATNEYGEVETSSKLLLOATPQFHPGYPLKEKYGAVGSTLRHLHVMYIGRPVPA 24162  
QY 238 ----- NVF- 240  
Db 24163 MTWFHGOKLLQNSENITIENTEYHLMVKMQKTHAGKYKQLNSNVEFTVDAILDVEI 24222  
QY 241 ----- 240  
Db 24223 QDKPKDPTGPIVIEALLKNSAVISMKPPADGGGSWITNTYVVEKCEAKEGAQWVSSAIS 24282  
QY 241 ----- 240  
Db 24283 VITCRIVNLTENAGYFRVSAQNTFGISDPLEVSSVVIKSPPEKPGPKPTITAVTKD 24342  
QY 241 SCSV ----- 244

Db 24343 SCVVAWPPASDGGAKIRNYILEKREKKONKWISVTTEEIRETVFSVKNLIEGLEVEFRV 24402  
QY 245 ----- 244  
Db 24403 KCENLGSESEWSEISEPITPKSDVPIQAPHFKBELNLNVRYSQSNATLVCVKVTHGPKPIV 24462  
QY 245 ----- 244  
Db 24463 KWVRQKEIADGLKYRIOBFKGGYHQLLIASVTDDATVYQVRATNQGSGVSGTASLEV 24522  
QY 245 ----- 244  
Db 24523 EYPAKIHLPKTTLEGMAVHALRGEVVSIKIPFSGKPDVITWQKGQDLIDNNGHYQIVIT 24582  
QY 245 ----- 244  
Db 24583 RSFTSLVPFNGVERKDAGFYVCAKNRFGIDQKTVELDVADVPDPGRGVKVSASRDSVN 24642  
QY 245 ----- 244  
Db 24643 LTWTEPASDGGSKITNYIVEKCATTAERLVRVQARETRYTVINLFGKTSYQFRVIAENK 24702  
QY 245 ----- 244  
Db 24703 FGLSKPSEPSEPTITTKEDKTRAMNYDEEVDREVSMTKASHSTKELYEKYMAEDLGR 24762  
QY 245 ----- MHEA- 248  
Db 24763 GEGFIVHRCVETSSKTYMAKFVKVKGTDQVLVKEISILNIAHRNHLHLHESFESMEE 24822  
QY 249 ----- LHNH- 252  
Db 24823 LYMWIFFISGLDIFERINTSAFELNEREIVSYVHVCEALQFLHSHNIGHFDIRPENIY 24882  
QY 253 ----- 252  
Db 24883 QTRRSSTIKIEFGQARQLKPGDNFRLFTAPEYAYEYVHQHDVSTATDMNSLGLTVVY 24942  
QY 253 ----- YT- 254  
Db 24943 LLSGINPFLAETNQIENIMNAEYTFDEAFKEISIEAMDVFDRLLVVERKSRMTASEA 25002  
QY 255 ----- QKSLSLS- 261  
Db 25003 LQHPWLKQKIERVSTKVIRTLKHRRYYHTLTKDLNVMVSAARISCGGAIRSKGYSVAK 25062  
QY 262 ----- 261  
Db 25063 VKVASTIEIGPVSQIMHAVGEGGHVKYVCKIENYDQSTQVWTWYFGVROLSENKEYEITY 25122  
QY 262 ----- 261  
Db 25123 EDGVAILYVKDITKLDGTYRCKVNVNDYGEDSSYAEFLVKGVREYDYVYCRRTMKIKRR 25182  
QY 262 ----- 261  
Db 25183 TDTMRLLERPBEFTLPYLNKTAAYGVENVRFGVTITVHPBPHVTWYKSGQIKPGDNKKY 25242  
QY 262 ----- 261  
Db 25243 TFSKGLYQLTINSVTDDAEYTVVARNKYGEDSCAKLTVTLHPPPTDSTLRPMFKR 25302  
QY 262 ----- 261  
Db 25303 LLANAEQEQSGVCFEIRVSGIPPTLKWEKDGQPLSLGNPIEIIHGLDYALHIRTDL 25362  
QY 262 ----- 261  
Db 25363 PEDTGYRVTATNTAGTSCQAHQLQVERLRYKKQEFKSEHERHVOKIDKTLRMAEIL 25422  
QY 262 ----- 261

Db 25423 SCTESVPLTQVAKAALREAAVLYKPAVSTKTVKGFRLEIEBKEERKLMPDYDPEPRK 25482  
QY 262 -----PGK 264  
Db 25483 YKOTTIEEDQRIKQVPMNDKMKWKKIRDOVEMPGK 25518  
RESULT 15  
Q8WB3  
ID Q8WB3 PRELIMINARY; PRT; 26926 AA.  
AC Q8WB3;  
DT 01-MAR-2002 (TremBLrel. 20, Created)  
DT 01-MAR-2002 (TremBLrel. 20, Last sequence update)  
DE 01-JUN-2002 (TremBLrel. 21, Last annotation update)  
DE N2B-titin isoform.  
GN TTN.  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
OX NCBI\_taxid=9606;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=20309627; PubMed=10850961;  
RA Freiburg A., Trombitas K., Hell W., Cazorla O., Fougereusse F.,  
RA Centner T., Kolmerer B., Witt C., Beckmann J.S., Gregorio C.C.,  
RA Granzier H., Labeit S.;  
RT "Series of exon-skipping events in the elastic spring region of titin  
RT as the structural basis for myofibrillar elastic diversity.";  
RL Circ. Res. 86:1114-1121(2000).  
RN [2]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=21573839; PubMed=11717165;  
RA Bang M.L., Centner T., Fornoff F., Geach A.J., Gotthardt M.,  
RA McNabb M., Witt C.C., Labeit D., Gregorio C.C., Granzier H.,  
RA Labeit S.;  
RT "The complete gene sequence of titin, expression of an unusual ~700  
RT kDa titin isoform and its interaction with obscurin identify a novel  
RT Z-line to I-band linking system.";  
RL Circ. Res. 89:1065-1072(2001).  
DR EMBL; AJ277892; CAD12455.1;  
DR InterPro; IPR000282; CytoK\_receptor\_2.  
DR InterPro; IPR000719; Euk\_pkinase.  
DR InterPro; IPR000577; FGGY\_kin.  
DR InterPro; IPR003961; FN\_III.  
DR InterPro; IPR001092; HLH\_basic.  
DR InterPro; IPR003599; Ig.  
DR InterPro; IPR003598; Ig\_c2.  
DR InterPro; IPR003600; Ig\_like.  
DR InterPro; IPR003006; Ig\_MHC.  
DR InterPro; IPR002016; Peroxidase.  
DR InterPro; IPR002290; Ser\_thr\_pkinase.  
DR InterPro; IPR001245; Tyr\_pkinase.  
DR Pfam; PF00041; fn3; 132.  
DR Pfam; PF00047; Ig; 91.  
DR Pfam; PF00069; pkinase; 1.  
DR ProDom; PD000001; Euk\_pkinase; 1.  
DR SMART; SM00060; FN3; 132.  
DR SMART; SM00409; IG; 111.  
DR SMART; SM00408; IGG2; 93.  
DR SMART; SM00410; IG\_like; 3.  
DR SMART; SM00220; S\_TKC; 1.  
DR SMART; SM00219; TykK; 1.  
DR PROSITE; PS00933; FGGY\_KINASES\_1; UNKNOWN\_1.  
DR PROSITE; PS00038; HELIX\_LOOP\_HELIX; UNKNOWN\_1.  
DR PROSITE; PS00290; IG\_MHC; UNKNOWN\_1.  
DR PROSITE; PS00435; PEROXIDASE\_1; UNKNOWN\_1.  
DR PROSITE; PS00011; PROTEIN\_KINASE\_DOM; 1.  
DR PROSITE; PS00109; PROTEIN\_KINASE\_TYR; UNKNOWN\_1.  
SQ SEQUENCE 26926 AA; 2993285 MW; 169AB42637A7C1FB CRC64;  
Query Match 33.9%; Score 485; DB 4; Length 26926;  
Best Local Similarity 0.7%; Pred. No. 0.0002;  
Matches 177; Conservative 44; Mismatches 42; Indels 23713; Gaps 55;

QY 2 WLK----- 4  
Db 1543 WLKNSDIIVPHKPKIRIEGKTGEAALKIDSTVSQDSAWYTATINKAGROTRTCKVNV 1602  
QY 5 -----AFYDK----- 9  
Db 1603 VFAEPEPERKLIIPRGTYRAKEIAAPELEPLHLRYGOEOWEGLDYDKERQKQKPPFKK 1662  
QY 10 ----- 9  
Db 1663 LLSLRKRFPAHFECRLTPIGDPTMVVWHLHDGKPLEAANRLRMINEFGYCSLDYGVAY 1722  
QY 10 ----- 9  
Db 1723 SRDSGIITCRATNKVGTDTHTSATLIVKDEKSILVESQLPEGRKGLQRIEELERMAHEGAL 1782  
QY 10 ----- 9  
Db 1783 TGVTTDOKEKOKPDIVLYPEPVRVLEGETARFCRVGTGPQKVNWYLNGLIRKSRFR 1842  
QY 10 ----- 9  
Db 1843 VRYDGIHYLDIVDCKSDTGEVKVTAENPEGVIEHKVKLEIQOQREDFRSLRRAPPEPRPE 1902  
QY 10 ----- 9  
Db 1903 FHVHEPGKIQFEVOKVDRPVDTTETKEVVKLRAERITHEKVPESBELRSKFKRRTBEG 1962  
QY 10 ----- 9  
Db 1963 YYEAITAVELKSRKDESYELLRRKTKDELHLHWTKEELKALAEKGKITITFTPKDK 2022  
QY 10 ----- 9  
Db 2023 IELSPSEAPKIFERIOSQTVGQSDAHFRVVRVVGKPDPECEWYKNGVKIERSRIYVW 2082  
QY 10 ----- 9  
Db 2083 PEDNVCVLIRDTAEDSASIMVKAINIAGETSSHAFLVQAKLIITFQELQDVVAK 2142  
QY 10 ----- 9  
Db 2143 DTMATFECETSEPFVKVYKDGMEVHEGDKYRHSORKVHFLSILTIDTSDAEDYSCVL 2202  
QY 10 ----- 9  
Db 2203 VEDENWTTAKLIVEGAVVEFKELQDIEVPEYSYSGELECIIVSPENIEGKWHNDVELKS 2262  
QY 10 ----- 9  
Db 2263 NGKYTITRRGRQNLTVKDVTKEDGGEYSFVIDGKTKTKLKKMPRIAILQGLSDQKVC 2322  
QY 10 ----- 9  
Db 2323 EGDIVQLVKVLSVESVEGVWMDQGEVQPSDRHVHIVDKOSHMLLIEDMTKEDAGNYSFT 2382  
QY 10 ----- 9  
Db 2383 IPALGLSTGRSVSYVDVITPLKDVNVIBGTRKAVLECKVSPDVTSVKWLNDQIKPD 2442  
QY 10 ----- 9  
Db 2443 DRVQAIKGTQKRLVINRTHASDEGPYKLIIVGRVETNCNLSVEKIKIIRGLDLTCTETQ 2502  
QY 10 ----- 9  
Db 2503 NVFVEVLSHGIDVLNFKDKETKPSKYKIEAHGKIYKLTVLNMMKDDGKTYTFYAGE 2562  
QY 10 ----- 9  
Db 2563 NITSGKLTVAGGAISKPLTDQTVAESQEAPEVCEVANPDSKGEWLRDGHPLPLTNIRSE 2622

QY 10 ----- 9  
Db 2623 SDGHRRLIIAANKLDDIGETTYKVATSKTSAKLKVEAVKIKKTLKLNLTVTETQDAVTV 2682  
QY 10 ----- 9  
Db 2683 ELTHPNVKGWIKNGVLESNEKYAISVKGTTIYSLRIKNAICVDESIVGFRGLGASA 2742  
QY 10 ----- 9  
Db 2743 RLHVEIVKIIKKPKDVTALENATVAFEVSVDHTVPVKHFKHVEIKPSDKHRLVSRKV 2802  
QY 10 ----- 9  
Db 2803 HKMLQNIQSPDAGEYAVVGQLECKAKLFEVTLHITKTMKNIETPETKTASFECEVSHF 2862  
QY 10 ----- 9  
Db 2863 NVPSMWLKNQVEIEMSEKFIYVQGLHQLIIMNTSTEDSAEYTFVCGNDQVSATLTVTP 2922  
QY 10 ----- 9  
Db 2923 IMITSMKLDINAEKDTIIFETVTNYEGISYKWLKNGVEIKSTDKCOMRTKKLTHSLNIR 2982  
QY 10 ----- 9  
Db 2983 NVHFGDAADYTFVAGKATSTATLYVEARHIEPRKHIDIKVLEKKRAMFECEVSEPDITV 3042  
QY 10 ----- 9  
Db 3043 QNMKDDQLOITDRIKIQKEKYVHRLIIPSTRMSDAGKTVVAGNVSTAKLFEVGRDVR 3102  
QY 10 ----- 9  
Db 3103 IRSIKKEVQVIEQRAVVEFEVNEDDVAHWYKDGIENFQVQERHKYVVERRIHRMFIS 3162  
QY 10 ----- 9  
Db 3163 ETRQSDAGEYTFVAGNRSSVTLVNAPEPPVQLQELPVTYVQSGKPARCAVISGRPOP 3222  
QY 10 ----- 9  
Db 3223 KISWYKEQLLSTGFKFLHDGQYTLLEIAPPEDAAVYTCERKNDYGVATTSASLV 3282  
QY 10 ----- 9  
Db 3283 EVPEVSPQOEMPVYPPAIITPLQDVTSEGOQPARFCRVSGTDLKVSWSKDKKIKPSR 3342  
QY 10 ----- 9  
Db 3343 FFRMTQFEDTYQLEIAEAYPEDEGTYTFVNASNAVQVSSSTANLSLEAPESILHERIEQEI 3402  
QY 10 ----- 9  
Db 3403 EMEMKFFSSFLSABEGLHSAELQSKINETLELLESVPYSTKFDSEKGTGPIFIKE 3462  
QY 10 ----- 9  
Db 3463 VSNADISMGDVATLSVTVIGIPKPKIQWFFNGVLLTPSADYKFVFDGDDHSLIILFTKLE 3522  
QY 10 ----- 9  
Db 3523 DEGEYTCMASNDYKGTICSAYLAKNSKGEGHKDTESAVAKSLEKLGCPPHFLKELK 3582  
QY 10 ----- 9  
Db 3583 PIRCAOGLPAIFEYTVVGEPAFTVTFWFKENKQLCTSVYVTI IHNPNGSGTFIVNDPQRED 3642  
QY 10 ----- 9  
Db 3643 SGLYICKAENMLGESTCAAEELLVLEDTMDTPCKAKSTPEAPDPQTPKGPAAVEAL 3702  
QY 10 ----- 9

Db 3703 DSQEIATFVKDTILKAALITEENQQLSYEHIAKANELSSQLPLGAQELQSILOQDKLTP 3762  
QY 10 ----- 9  
Db 3763 ESTREFLCINGSIHFOPLKEPSNLOLQIVSOKTFSKESGILMPEEPETOAVLSDTEKIF 3822  
QY 10 ----- 12  
Db 3823 PSAMSTEQINSLTVEPLKTLAEPEGNYPOSSIEPPMHSHSYLTVAEVLSPKERTVSDTN 3882  
QY 13 ----- 12  
Db 3883 REQRTVLQEQEAQSALILSQSLAEGHVESLQSPDMISQVNYBPLVPSEHSCTEGGKILI 3942  
QY 13 ----- 12  
Db 3943 ESANPLENAGQDSAVRIEKGSLRFFPLALEEKQVLLKEHSDNVVMPDQIIIESKREPVA 4002  
QY 13 ----- 12  
Db 4003 IKKQVEQGRDILLSKESLISGIPPEQRLNLKIQICRALQAAVASEQPGLESEWLNIEKV 4062  
QY 13 ----- 12  
Db 4063 EVEAVNITQEPRHIMCMLVTSKSVTEEVTHIIEVDVDPQMANLKMELRDALCAIIVEEI 4122  
QY 13 ----- 12  
Db 4123 DILTAECPRIQOGAKTSLQEMDSFSGSQKVEPITEPEVESKYLIISPEEVSYNFVOSRVK 4182  
QY 13 ----- 17  
Db 4183 YLDATPVTGVSASVDEKQDESLEKPSKEESESSESTEEVATVKIQEAGSGFIKEDG 4242  
QY 18 ----- 17  
Db 4243 PMIHTPLVDTVSEGDIVHLTTSITNAKEVNWYFENKLVPSDEKFKCLQDQNTYTLVIDK 4302  
QY 18 ----- 17  
Db 4303 VNTEHQGEYVCEALNDSGKTATSAKLTVVYKRAAPVIKRKIEPLEVALGHAKFTCEIQS 4362  
QY 18 ----- 23  
Db 4363 APNVRQWFKAGRIEYESDKSIRSSKYISSLEILLTQVVDCGEYTCKASNEYGSVCTA 4422  
QY 24 ----- 23  
Db 4423 TLTVTVPGGKKYRKLLPERKPEPEEVLKSVLKRPEEPEKPEKLEKVKKPAVPE 4482  
QY 24 ----- 23  
Db 4483 PPPPKPEEVEVPTVTKRERKIPEPTKVPEIKPAIPLPAPEPKPKEAEVTKIKPPPEP 4542  
QY 24 ----- 23  
Db 4543 EPTPIAAPVTVPVVGKKAERAKPEEAKPKGPIKGVPKKTPSPIERERKLRPGSGEK 4602  
QY 24 ----- 23  
Db 4603 PPDEAPFTYQKAVPLKFKVKEIKDIIILTESEFVGSAAIFELVSPSTAITTWMKDGSNIR 4662  
QY 24 ----- 23  
Db 4663 ESPKHRFIADGDKRKLHIIDVQLSDAGEYTCVLRGLGNKKTSTAKLVEELPVRVFKTLE 4722  
QY 24 ----- 23  
Db 4723 EEVTVVKGOPYLSCELNKERDYYVWRKDGKIVVEKPRIVPGVIGLMRALTIINDADDTDA 4782  
QY 24 ----- 23

Db 4783 GTYTVTVENANNLECSVCVVEVIRDLVLPKPIRDQHVKPKGTAFACDIAKDTNPKWF 4842  
QY 24 ----- 23  
Db 4843 KGYDEIPAPNDKTEILRDGNHLYLKINAMPEDIAEYAVEIEGKRYPAKLTGLGERVEL 4902  
QY 24 ----- 23  
Db 4903 LKPIEDVTIYEKESAFDAEISEADIPGQWKLGLLRLSPPTCEIKAEGRFLRLRVK 4962  
QY 24 ----- 23  
Db 4963 LDQAGEVLYQALNATITALLTVKEIELDFAVPLKDVTPERRQARFECVLTREANVWSK 5022  
QY 24 ----- 23  
Db 5023 GPDIIKSSDKFDIADGKKHILVINDSQDDEGVYTAEVGKKTSARLFVIGIRLKFMPSP 5082  
QY 24 ----- 23  
Db 5083 LEDQTVKEGETATFVCELSHEKMHVWFKNDAKLHTRSRTLISSEKTHKLEMKVTLDD 5142  
QY 24 ----- FDKVAE----- 30  
Db 5143 ISQIKAOVKELSSTAQLKVLEADPYTVKLHDKTAVEKDEITLKEVSKDVPVKWFKDGE 5202  
QY 31 ----- 30  
Db 5203 EIVSPKYSIKADGLRILIKKADLKDGEYVDCGTDKTKANTVVEARLIKVEKPLYG 5262  
QY 31 ----- 30  
Db 5263 VEVFVGETAHFEIELSEPDVHGOWKLGOPLTASPDCEIIEDGKKHILHNCQLGTMGE 5322  
QY 31 ----- KLKE----- 34  
Db 5323 VSFOAANKSAANLVKELPLIFITPLSDVKVFKEKAEKFECEVSREPKTFWLKGTQEI 5382  
QY 35 ----- AFMD----- 38  
Db 5383 TGDDRFLIKDGTKHSWIKSAAPDEAKYFPAEDKHTSGKLIIEGIRLKLPLKDV 5442  
QY 39 ----- 38  
Db 5443 AKEKESAVTVELSHDNIRVKFNQDRLHTRTSVSMODEGKTHSITEKDLSDIDTSQIR 5502  
QY 39 ----- 38  
Db 5503 VEMGMSSEAKLTVLEGDPYFTGKLQDYGVEKDEVILOCEISKADAPVKWFKDGEIKP 5562  
QY 39 ----- 38  
Db 5563 SKNAVIKADGKRMLILKALKSDIGQVTCDCGTDKTSGLDIEDREIKLVRPLHSVEVM 5622  
QY 39 ----- 38  
Db 5623 EFTARFETEISEDIIHANWKLKGEALLQTPDCEIKEEGKIHSVLHNCRLDQGTGGVDFQ 5682  
QY 39 ----- 38  
Db 5683 RANVSSAHLRVKPRVIGLLRPLKDVTVTAGETATFDCELSYEDIPVEWYLGKKLEPSD 5742  
QY 39 ----- 38  
Db 5743 KVPRESGVHTLRLRDVKLEDAGEVOLTAQDFKTHANLFVKEPPVEFTKPLEDQTVBEG 5802  
QY 39 ----- KTHTC----- 43  
Db 5803 ATAVLECEVSRENKVKWFKNGTEILSKSYEIVADGVRKLVTHDCTPEDIKTYTCDAK 5862  
QY 44 ----- 43  
Db 5863 DFKTSCNLNVPPHVEFLRPLTDLOVREKEMARFECELSRENKVKWFKDGAETKKGKY 5922

QY 44 ----- 43  
Db 5923 DIISKGAVRILVINKCLLDDEAEYSCVETARTSGMLTVLEBEAVFTKNLANIEYSETDT 5982  
QY 44 ----- 43  
Db 5983 IKLVCEVSPGAEVIWYKGDDEIIETGRYEILTEGRKRILVIONAHLEDAGNYNCRPLSS 6042  
QY 44 ----- 43  
Db 6043 RTDGKVKVHELAAEFISKPONLEILEGEKAEFVCSISKESFPVQWKRODKTLESCKYDV 6102  
QY 44 ----- 43  
Db 6103 IADGKKRVLVKDATLQDMGTYYVMVGAAARAAHLTVIEKLRIVVPLKDRVKEQOEUVF 6162  
QY 44 ----- 43  
Db 6163 NCEVTEGAKAKWFRNEEAI FDSKYYIILOKDLVYTLRIRDAHLDDQANYNVSLTNHRGE 6222  
QY 44 ----- 43  
Db 6223 NVKSAANLIVEEEDLRIVEPLKDIETMEKKSVTFWCKVYNRLNVLTKTKNGEVPFDNRV 6282  
QY 44 ----- 43  
Db 6283 SYRVDKVKHMLTIKDCGFPDEGEYIVTAGQDKSVAELLIIEAPTEFVEHLEDQTVTEFDD 6342  
QY 44 ----- 43  
Db 6343 AVFSQLSREKANVWYRNGREIKGKKYKPEKDGSIHRLIHKDRLDDECEYACGVEDR 6402  
QY 44 ----- 43  
Db 6403 KSRARLFVEETPVEIIRPPQDILEAPGADVFLAELNKKDKVEQVLRNMNMVVQDKKHOM 6462  
QY 44 ----- 43  
Db 6463 MSEGKIHLQICDIKPRDQGEYRFTAKDKEARAKLEAAAPKIKTADQDLVVDVGKPLTM 6522  
QY 44 ----- 43  
Db 6523 VVPYDAYPKABEAFKENEPLSTKTIDTBTABOTSPRILEAKKGDKGRYKIVLONKHGAE 6582  
QY 44 ----- 43  
Db 6583 GFINKLVIDVPGVPVRLNLEVTETFDGEVSLAWEEPLTDGSGKLIIGYVVERDRIKRTWVLA 6642  
QY 44 ----- 43  
Db 6643 TDRAESCEFTVTGLQKGGVEYLFVRSARNRVGTGEPVETDNPVEARSKYDVPGLNVTI 6702  
QY 44 ----- 43  
Db 6703 TDVNRFGVSLFWPEPYDGGAEITNYVIELRDKTSIRMDTAMTVRAEDLSATVTDVVEGQ 6762  
QY 44 ----- 43  
Db 6763 EYSFRVRAQNRIGVKGKPSAATPFVKVADPIERPSPVNLTSDDQTSQSSVLKWPPLKDG 6822  
QY 44 ----- 43  
Db 6823 GSPILGYIIERCCEGKNWIRCNMKNLVPCLTYKVTGLEKGNKLYLRVSAENKAGVSDPSE 6882  
QY 44 ----- 43  
Db 6883 ILGPLTADDAFVEPTMDLSAFKDGLEIVPNPITILPSTGYPRPTATWCFGDKVLETFD 6942  
QY 44 ----- 43  
Db 6943 RVKMKTL SAYAELVISPERSDKGIYTLKLENRVKTIISGEIDVNVVIARPSAPKELKFGDI 7002



QY 44 ----- 43  
Db 7003 TKDSVHLTWEPDDDDGSLTYGVVEKREVSRTKTKVDMFVTDLEFTVPDLVQOGKEYLF 7062  
QY 44 ----- 45  
Db 7063 KVCARNKCGGEPAYVDEPVNMSTPATVPPDENVKWRDRTANSIFLTWDPKNDGGSRI 7122  
QY 46 ----- 47  
Db 7123 KGIVVERCPGSKWVACGEPVAETKMEVTGLEEGKWYAYRVKALNRQASKPSRTEI 7182  
QY 48 ----- 52  
Db 7183 QAVDTQEAPEIFLDVLLAGLTVKAGTKIELPATVTGKPEKITTKADMILKQDKRITI 7242  
QY 53 ----- 58  
Db 7243 ENVPKSTVTI VDSKRSDTGTIIEAVNVCGRATAVEVNVLDKPGPPAAFDITDTVTNES 7302  
QY 59 ----- 58  
Db 7303 CLLTWNPDRDDGSKITNVVERRATDSEVHKLSTVKTDFKATKLIPNKEYIFRVA 7362  
QY 59 ----- 58  
Db 7363 ENMYGVGEVQASPI TAKYQFDPGPPPTRLPESDITKDAVTLTWCEPDDDDGGSPITGYV 7422  
QY 59 ----- 58  
Db 7423 ERLDPTDKWRCNMPVDTTYRVKGLTNKKYFRVLAENLAGPKPSKSTEPILLIKD 7482  
QY 59 ----- 58  
Db 7483 PIDPPWPPGKPTVKDVGKTSVRLNWTKEPHDGGAKTESVVIEMLKTGTDEWVRVAEGVPT 7542  
QY 59 ----- 64  
Db 7543 TQHLPLGLEGGQYSFRVAVNKAAGESEPSPDPVLCREKLYPPSPRWLEVINITKNT 7602  
QY 65 ----- 73  
Db 7603 ADLKWTVPKDDGSPITNYIVEKRDVRRKGWQTVDTVTVDKTKCTVTPLEGSLYFRVAA 7662  
QY 74 ----- 83  
Db 7663 ENAIGSDYTEIEDSVLAKDTFTTGPCPYALAVVDVTKRHRVDLKWEPPKNDGGRPIQRV 7722  
QY 84 ----- SHE 86  
Db 7723 IEKKERLGRWVKAGKTAGDCNFRVTDVIEGTEVQFQVRAENEAGVGHPSPTIELSIE 7782  
QY 87 DP----- 88  
Db 7783 DPTSPSPPLDLHVTDAGRKHIAIAWKPKPKNGGSPIIGYHVEMCPVGTCKWVRVSRPI 7842  
QY 89 -EVKF----- 92  
Db 7843 KDLKFVEGVDPKEYVILRVRAVNAIGVSEPSEISENVVAKDPCKPTIDLETHDIYI 7902  
QY 93 ----- 98  
Db 7903 EGEKLSIPFRAVPPTVSWHKDGEVKASDRLTMKNDHISAHLEVPKSVRADAGIYTI 7962  
QY 99 ----- 98  
Db 7963 TLENKLGASASINVKVIGLPGCKDIKASDITKSSCKLTWEPPEPDGPTILHYVLERR 8022  
QY 99 ----- 98  
Db 8023 EAGRRTYIPVMSGENKLSVTKDLPNGEYFFRVKAVNVKGGEYIELKNPVIAQDPKOP 8082  
QY 99 ----- 104  
Db -----VEVHNA-----

Db 8083 POPPVDEVHNPTAEAMTITWKPPLYDGGSKIMGYIIIEKIAKEERWKRCHLHVILTY 8142  
QY 105 ----- 104  
Db 8143 TAKGLEEGKEYOFRVRAENAAIGSEPSRATPPTKAVDPIADPKVILRTSLEVRKGDEIAL 8202  
QY 105 ----- 104  
Db 8203 DASISGSPYTTITWIKDENVIYVPEEIKKRAAPLVRRRKGEVQEEFVLPLOTQRLSDNS 8262  
QY 105 ----- 104  
Db 8263 KKGESQLRVDRSLRPDHLGYMIKVENDHGIAKAPCTSVLDTGPPPINFVFEIDIRKTSVL 8322  
QY 105 ----- 108  
Db 8323 CKWEPPLDGSGSEIINYLEKKDKTKRPSSEWIVVTLRHCKYSVTKLIEGKEYLFRVA 8382  
QY 109 ----- 108  
Db 8383 ENRFGPCPCVSKPLVAKDPFGPPDAPDKPIVEDVTSNSMLVKWNEPKONGSPILGYWLE 8442  
QY 109 ----- 108  
Db 8443 KREVNSTHWSRVNKSLLNALKANVDGLLEGLTYVFRVCAENAAAGPGKFSPPDPKTAHDP 8502  
QY 109 ----- 108  
Db 8503 ISPPGPIPRVTDTSSTTIELEWEPAPNGGGEIVGYFVDKQLVGTNEWSRCTERMIKVR 8562  
QY 109 ----- 108  
Db 8563 QYTVKIREGADYKLRVSANNAAGEPPGETQPVTVAEPOEPAPVELDVSKGQIIMAG 8622  
QY 109 ----- 108  
Db 8623 KTLRIPAVVTGRVPTKVMTKEGELDKDRVIDNVGTSKSELIKDALRKDHRYVITAT 8682  
QY 109 ----- 108  
Db 8683 NSCGSKFAAAARVEVDFVPGFVLDLKPVVTNRKMLLNWSDPEDDGGSEITGFIERKDAK 8742  
QY 109 ----- 108  
Db 8743 MHTWROPIETERSKCDITGLLEGQEKYFRVIAKNKFCGCPPEIGPILAVDPLGPPTSPE 8802  
QY 109 ----- 108  
Db 8803 RLTYTERTKSTITLDWKEPRNGSGSPIQGYIEKRRHDKPDPFERNKRLCPTTSFLVENL 8862  
QY 109 ----- 108  
Db 8863 DEHQMYEFRVAVNEIGSESESLPNVVIQDDVEVPPTIKRLSVRGDTIKVKAGEPVHIP 8922  
QY 109 ----- 108  
Db 8923 ADVTGLPMPKIEWSKNETVIEKPTDALQITKEEVSSEAKTELSIPKAVREDKGTVTVA 8982  
QY 109 ----- 108  
Db 8983 SNRLGVSFRNVHVEVYDRPSPRNLAVTDIKAESCYLTWDAPLDNGGSEITHVIDKRDA 9042  
QY 109 ----- 108  
Db 9043 SRKKAEEVNTAVAKRYGIMKLIPINGQYEFVRVAVNKGYSDECKSOKVVIQDPYRLP 9102  
QY 109 ----- 108  
Db 9103 GPPGPKVILARTKGSMLVSWTPPLDNGGSPITGYWLEKREEGSPYWSRVSRAPITKVGLK 9162  
QY 109 ----- 108  
Db -----

Db 9163 GVEFNVPRLLGKGYQFRAMAINAAGIGPPSEPSDEPVAGDPIFPFPPSPCEVVKDTKS 9222  
QY 109 ----- 108  
Db 9223 SISLGWKPAPKGGSPKIGYIVEMOEBGTTDMKRVNEBDKLTTCCECVVNLKELRYRF 9282  
QY 109 ----- 108  
Db 9283 RVKAVNEAGESEPSDTTGEIPATDIOEBPEVFDIGAQDCLVCKAGSQRIRPAVIRGPT 9342  
QY 109 ----- 108  
Db 9343 PKSWFEDGKAKKAMKGVHDIPEDAQLETAENSSVIIPECKRSHTGYSITAKNKAQ 9402  
QY 109 ----- 108  
Db 9403 KTANCRVKVMYVPPKDLKYSDIIRGSCRLSWKMPDDDDGDRIRKGYVIEKRTIDGKAWT 9462  
QY 109 -----REEQY----- 113  
Db 9463 KVNPCGSTTFVVPDLLSEQQYFFRVRRAENRFGIGPPVETIQTARTDPIVPPDPPIKLK 9522  
QY 114 ----- 113  
Db 9523 IGLITKNTVHLSWPKKNDGGSPVTHYIVECLANDPTGTKEAWRQCNKRDVEELQFTVE 9582  
QY 114 ----- 113  
Db 9583 DLVEGGEYFRVKAVNAAGVSKPSATVGPCQCRDMPSPIDLKEFMEVEBGTNNIVAK 9642  
QY 114 ----- 113  
Db 9643 IKGVPFPTLTFKAPPKPNKEPVLVDTHVKNLVDDTCTLVIPOSSRSDGLYTITAV 9702  
QY 114 ----- 113  
Db 9703 NNIGTASKEMNLVLRGPPVGPPIKFESVSADQMTLSWFPPKDDGGSKITNYVIEKREA 9762  
QY 114 ----- 113  
Db 9763 NRKTVWVSSPEKCTYTIKPLEGHEYFRIMAQNKYGIGEPDSEPETARNLFSVPGA 9822  
QY 114 ----- 113  
Db 9823 PDKPTVSSVTRNSMTVNWEPEYDGGSPVTGYWLEMKDTSKRWRKVRNRPDIKAMTLGVS 9882  
QY 114 ----- 113  
Db 9883 KYVTGLIEGSDYQFRYAINAAGVPASLSPDAPATARDPIAPPGPPPKVTDWTKSSADL 9942  
QY 114 ----- 113  
Db 9943 EMSPLKDGSKVTGYIVEYKEEGKEEKGKDEVRGTLVVTKLEGAFFKFRVSAVN 10002  
QY 114 ----- 113  
Db 10003 IAGIGEPGEVDVIEKDRLVSPDLQDASVRDRIVHVAGGVIIRIIAYVSKPPPTWTN 10062  
QY 114 ----- 113  
Db 10063 MNERTLPOEATIETTAISSMWIKNCQRSHOGVYSLAKNEAGERKRTIIVDVLDPGPV 10122  
QY 114 ----- 113  
Db 10123 GTPFLAHLNLTWFSPEDDGGSPITNVVIEKRESDRRAWTPVTYTYTRQNAIVQG 10182  
QY 114 ----- 113  
Db 10183 LIQKAYEFFRIAAENSIGMPFVETSEALVIREPITVPERPEDLEKVEKNTVTLTWNP 10242  
QY 114 ----- 113  
Db 10243 PKYDGGSEIINVLRESRLIGTEKFKHVTNDNLLSRKTYTVKGLKEGDTVEYRVSAVNIVGQ 10302

QY 114 ----- 113  
Db 10303 GKPSFCTKPIITCKDELAPPTLHLDPRDKLTIRVGEAFALTGRYSCKPKPKVSWFKDEADV 10362  
QY 114 -----NST----- 116  
Db 10363 LEDDRTHIKTTPATLALAEKIKAKRSDSGKYCVVVENSTGSRKGFQVNVWRPDPVGPV 10422  
QY 117 ----- 116  
Db 10423 SFDEVTKDYMVISWKPPDLDDGGSKITNYIIIEKKEVGKDVWMPVTSASAKTTCKVSKLLEG 10482  
QY 117 ----- 116  
Db 10483 KDYIFRIHAENLYGISDPLVSDSMKAKDRFRVPDAPDQPIVTEVTKDSALVTWKNKPHDGG 10542  
QY 117 ----- 116  
Db 10543 KPITNYILEKRETMKRWARTKDPHPYTKFRVPDLLLEGQCYEFRVSAENEIGIGDPS 10602  
QY 117 ----- 116  
Db 10603 PSKPVFAKDPIAKSPVPVNEAIDTTCNSVDLTWQPPRHGGSKILGYIYEVQKVGDEEM 10662  
QY 117 -----YRVSV----- 122  
Db 10663 RRANHTPESCPETKYKVTGLRDGQTYKFRVLAVNAAGESDPAHVPEPVLVKDRLEPPELI 10722  
QY 123 ----- 122  
Db 10723 LDANMAREQHIKVGDTLRLSAIIKGVPPKVTWKEDRDAPTAKRIDVTPVGSKLEIRNA 10782  
QY 123 -----LTV----- 125  
Db 10783 AHEDGGIYSLTVENPAGSKTVSVKVLVDKPGPPRDLEVSEIRKDCSYLTWKEPLDDGGS 10842  
QY 126 ----- 125  
Db 10843 VITNVVVERRDVASAQWSPLSATSATSKKSHFAKHLNEGNOYLFRVAAENQYGRGPEVETPK 10902  
QY 126 -----LH-----QDWL-- 131  
Db 10903 PIKALDPLHPGPPKDLHVDVDKTEVSLVWKNPDRGGSPITGYLYVEQBEGTQDWIKF 10962  
QY 132 ----- 131  
Db 10963 KTVNLECVVTGLOQGTKYFRVRKAENIVGLPDTTPIECQEKLVPPSVELDVKLIEG 11022  
QY 132 ----- 131  
Db 11023 LVVKAGTTVRFPATIRGVVPVPTAKWTTDGSSEIKTDEHYTVETDNFSSVLTIKNCLRRDGT 11082  
QY 132 ----- 131  
Db 11083 EYQITVSNAGSKTAVHLTVLDPGPPGTGINILDVTPHEMTISWQPKDDGGSPVINY 11142  
QY 132 ----- 131  
Db 11143 IVEKQDTRKDTWGVVSSGSSKTLKIPHLQKGEYFRVRAENKIGVGPPLDSTPTVAKH 11202  
QY 132 ----- 131  
Db 11203 KFSPPSPGKPVVTDITENAAATVSWTLPKSDGGSPITGYMYERREVTGKWRVKNKPLAD 11262  
QY 132 ----- 131  
Db 11263 LKFRVTGLYEGNTYFRVFAENLAGLSKPSDSDPIKACRPDKPPGPINPKLKDKSRET 11322  
QY 132 ----- 131  
Db 11323 ADLVWTKPLSDGGSPILGYVVECOKPGTAQWNRINKDELIRQCAFRVPGLLEGNEYRPI 11382

QY 132 ----- 131  
Db 11383 KAANIVGECEPRELAESVIAKDILHPPEVELDVTCDVITVRVGTIRILARVKGRPEPD 11442  
QY 132 ----- 131  
Db 11443 ITWTEGKVLVREKRVDLIQDLPRVELOIKEAVRADHGKIISAKNSSGHAQSAIVNVL 11502  
QY 132 ----- 131  
Db 11503 DRPGCONLKVNTVKENCTISWENPLDNGSGSEITNFIVEYRKPQKGHSIVASDVTKRL 11562  
QY 132 ----- 131  
Db 11563 IKANLIANNEYFRVCAENKVGVPITETKPTILAINPIDRGEPENLHIADKGKTFVYL 11622  
QY 132 ----- 131  
Db 11623 KWRRPDYDGGSPNLSYHVERRLKGDDWVERVHKSIKETHYMWDRCVENQIYEFVRVQTKN 11682  
QY 132 ----- 131  
Db 11683 EGGESDWKTEEVVYKEDLQKPVLDLKLGLSVLTVKAGDTIRLEAGYRGPFPPEVAWTKDK 11742  
QY 132 ----- 131  
Db 11743 DATDLTRSPRVKIDTRADSKFSLTKAKRSDGGKYVVTATNTAGSEFVAVATNVLDKPGP 11802  
QY 132 ----- 131  
Db 11803 VRNLKIVDVSSDRCTVCWDPDPEDDGCETONYLEKCEKRMVMWSTYSATVLTPTGTVTR 11862  
QY 132 ----- 131  
Db 11863 LIEGNEYIFRVAENKIGTGPTESKPVIATKYDKPGRPDPPEYTKVSKEEYVWVWNP 11922  
QY 132 ----- 131  
Db 11923 EYDGGKSITGYFLEKKEKHSTRWPVNVKSAIPERRMKVQNLDPDHEYQFRVAENEIGT 11982  
QY 132 ----- 131  
Db 11983 EPSLPSRPVAKDPIEPGPPPTNFRVVDTKHSITLGGWKPVYDGGAPIIIGYVEMRPKI 12042  
QY 132 ----- 131  
Db 12043 ADASPDGKRCNAAQLYRKEFTVTSLDENQEYEFVCAQNOVGIGRPAELKEAIKPE 12102  
QY 132 ----- 131  
Db 12103 ILEPPEIDLDASMRKLIIVRAGCPIRLFAIVRGRPAKVTWRKVIDNVVRKGQVDLVDT 12162  
QY 132 ----- 131  
Db 12163 MAFLVTPNSTRDDSGKSYTLVNPAGEKAVFVNVRLDTPGPVSDLKVSDVTKTSCHVSW 12222  
QY 132 ----- 131  
Db 12223 APPENDGSGQVTHYIVEKREADRKTWSTVTPVKKTSFHTNLVPGNEYFRVTAVNEYG 12282  
QY 132 ----- 131  
Db 12283 PGVPTDVPKPVLASDPLSEDPDRKLEVTMTKNSATLAWLPPLRDGAKIDGYITSYRE 12342  
QY 132 ----- 131  
Db 12343 EQPADRWTEYSVVKDLVLVVTGLKGGKKYKFRVAARNAGVSLPREAGVYEAKEQLLP 12402  
QY 132 ----- 131  
Db 12403 PKILMEQITIRAGKRLREAHVYGKPHPTCKWKGEDEVTTSSHLAVHKADSSILIK 12462  
QY 132 ----- 131

Db 12463 DVTRKDSGYISLTAENSSGTDQKIKVVVMDAPGPPQPPFDISDIDADACSLSWHIPLD 12522  
QY 132 ----- 131  
Db 12523 GGSNITNYIVEKCDVSRGDWVTALASVTYTSRVGKLIPGQEYIFRVAENRFGISEPLT 12582  
QY 132 ----- 131  
Db 12583 SPKMAQFPFGVPSEPKNARVTKVNKDCIFVAWDRPDSGGSPILIGYLIERKERNLLW 12642  
QY 132 ----- 131  
Db 12643 KANDTLVRSTPYCAGLVEGLEYSFRIYALNKAGSSPPSKPTEYVTARMPVDPGKPEVI 12702  
QY 132 ----- 131  
Db 12703 DVTKSTVSLIWARPKHDGSKIIIGYFVEACKLPDGKWVRCNTAPHQIPOEYATATGLEK 12762  
QY 132 ----- 131  
Db 12763 AOYQFRAIARTAVNISPPSESDPVTILAENVPPRIDLSVAMKSLTLVKAGTNVCLDATV 12822  
QY 132 ----- 131  
Db 12823 FGKPMPTVSWKKDGTLLKPAEGIKMAMQRLCTLEFSVNRKDSGDTTITAENSSGSKSA 12882  
QY 132 ----- 131  
Db 12883 TIKLKVLDKPGPPASVKINKMYSDRAMLSWEPPLEDGGSEITNYIVDKRETSRPNWAQVS 12942  
QY 132 ----- 131  
Db 12943 ATVPITSCSEKLEIEGHEYQFRICAENKYGVDPVFTEPAIAKNPYDPGRCDDPPVISNI 13002  
QY 132 ----- 131  
Db 13003 TRDHMTVSNKPADGGSPITGYLLEKRETOAVNWTKVNRKPIIERTLAKTLOQEGTEYE 13062  
QY 132 ----- 131  
Db 13063 FRVTAINKAGCPKPSDASKAAAYARDPOYPPAPPAPFKVYDTRSSVSLSWGKPAYDGGSP 13122  
QY 132 ----- 131  
Db 13123 IIGLYLVKRADSDNWRCNLPLONLQKTRFEVTGLMEDTQYQFRVAVNKGIGYSDPSDVP 13182  
QY 132 ----- 131  
Db 13183 DKHYPKDILIPPEGELDADLRKTLILRAGVTMRLYVPVKGRRPPPKITWSKPNVNLDRIG 13242  
QY 132 ----- 131  
Db 13243 LDIKSTDFDTFLRCENNVNKYDAGKYILTLENSCGKKEYTIVVKVLDTPGPPVNVTVKEIS 13302  
QY 132 ----- 131  
Db 13303 KDSAYVTWEPPIIDGGSPIINYVVKRDAERKSWSTVTTECSKTSFRVANLEEGKSYFFR 13362  
QY 132 ----- 131  
Db 13363 VFAENEYIGDPGETRDVAKASQTPGVVDLKVRSVSKSSCSGTGWKKPHSDGSGRIIGYV 13422  
QY 132 ----- 136  
Db 13423 VDFLTEENKQWRVMKSLSLQYSAKDLTEGKEYTFRVSAENENGEGTPEITVVVARDVVA 13482  
QY 137 ----- 136  
Db 13483 PDLDLKGLPDLCYLAKENSFRLKIPKGPAPSVWKKGEDPLATDTRVSESSAVNTT 13542  
QY 137 ----- 136

Db 13543 LIVYDQKSDAGKYTITLKNVAGTKBGTISIKVVGKPGIPTGPIKFDDEVTAEMTLKWP 13602  
Qy 137 ----- 136  
Db 13603 PKDDGSEITNYILLEKRDVNKWTCSAVQKTFTRVRLHEGMEYTFRVAENKYGVG 13662  
Qy 137 ----- 136  
Db 13663 EGLKSEPIVARHPFDVDPADPPPNIVDVHRDVSLSLTWDPKKTGGSPITGYHLEFKERNS 13722  
Qy 137 ----- 136  
Db 13723 LLWKRANKTPIMRDKFKVTGLTEGLEVEFRVMAINLAGVGKPSLPSEPVVALDPIDPPGK 13782  
Qy 137 ----- 136  
Db 13783 PEVINITRNSVTLIWTPEPKYDGGHKLTCYIVEKRDLPKSHMKANHVNPCEAFTVTDLV 13842  
Qy 137 ----- 136  
Db 13843 EGGKYEFRIRAKNTAGAISAPSESTETIICKDEYEAPTIVLDPTIKDGLTIKAGDTIVLN 13902  
Qy 137 ----- 136  
Db 13903 AISILGKPLPKSWSKAGKDIRPSDITQITSTPTSSMLTIKYATRKDAGEYTIATNPF 13962  
Qy 137 ----- 136  
Db 13963 TKVEHVKVTVLDVPCPGPVEISNVAEKATLTWTPLEDGSGPIKSYILEKRETSRLW 14022  
Qy 137 ----- 136  
Db 14023 TWSIEDIQSCRHVATKLIQNEGYFRVSAVNHYGKGPVQSEPKVMDRFPGPPGPEKPE 14082  
Qy 137 ----- 142  
Db 14083 VSNVTKNTATVSWKRPVDDGSEITGYHVERREKSLRWRAIKTPVSDLRCKVTGLQEG 14142  
Qy 143 ----- 142  
Db 14143 STYEFVSAENRAGIGPPSEASDSVLMKDAAYPGPPSNPHVTDTTKKSASLANGKPHYD 14202  
Qy 143 ----- 142  
Db 14203 GGLEITGVVEHQVKGDEAWIKDTGTALRITQFVVPDLOTKEKYNFRISAINDAGVGP 14262  
Qy 143 ----- 148  
Db 14263 AVIPDVEIVEREMAPDFELDAELRRTLVRAGLSIRIFVPKGRPAPEVVTWKDNLNKN 14322  
Qy 149 ----- 148  
Db 14323 RANIENTESFLLIIPECNRYDTGKFVMTTENPAGKSGFVNVVRVLDTPGPVNLNRPDI 14382  
Qy 149 ----- 151  
Db 14383 TKDSVTLHWDLPLIDGGSRTITNYIVEKREATKSYSTATTKCHKTKYKVTGLSGBCEYFF 14442  
Qy 152 ----- 151  
Db 14443 RVMAENEYGIGEPTETPEPVKASEAPSPDLSNIMDITKSTVSLAMPKPKHDGSKITGY 14502  
Qy 152 ----- 151  
Db 14503 VIEAQKGSQDWHITTVKGLECVVRNLTEGEYTFQVMAVNSAGRSAPRESRPVIVKEQ 14562  
Qy 152 ----- 151  
Db 14563 TMLPELDLGIYQKLVIAKAGDNIKEIVPLGRPKPTVTKWKGDIKQIQRVNFETTAT 14622  
Qy 152 ----- 151  
Db 14623 STILNINECVRSDSGPYPLFARNIVGEVDVITIQVHDIQPGPTGPIKFDEVSSDFVTF 14682

Qy 152 ----- 151  
Db 14683 WDPENDGGVPISNVVVMEMQTDSTTWELATTIVRTTYKATRLTTGLEQFRVKAQNR 14742  
Qy 152 ----- 151  
Db 14743 GVGPGITSACIVANYPFKPGPGTPOVTAVKDSMTISWHEPLSDGSGPILGYHVERKE 14802  
Qy 152 ----- 156  
Db 14803 RNLGILMOTVSKALVPGNIFKSSGLTDGIAYEFRVIAENMACKSPKSPSEPMALADPIDP 14862  
Qy 157 ----- 156  
Db 14863 PGKVPNLNITRHTVTLKWKAPETGTFKITSYIVEKRDLPNGRWLKNFNSILENEFTVS 14922  
Qy 157 ----- 156  
Db 14923 GLTEDAAVEFRVIAKNAAGAISSPPSPSDAITCRDDVEAPKIKVDVKFKDTVILKAGEAF 14982  
Qy 157 ----- 156  
Db 14983 RLEADVSGRPPPTMWSKDGKELEGTAKEIKIADFSTNLVKNKSTRDRSGAYTLTATNP 15042  
Qy 157 ----- 156  
Db 15043 GGFAKHIFNVKVLDRPGPPGPEGLAVTEVSEKCVLSWFPPLDDGGAKIDHIVIVOKRETSR 15102  
Qy 157 ----- 156  
Db 15103 LAMTNVASEQVTKLVTKLLKGNEYIFRVMAVKNYGVGPESEPVLA VNPYGPDPKP 15162  
Qy 157 ----- 156  
Db 15163 NPEVTTITKDSMVVWCWGHDPDSDGSEIINIVERRRKAGORWIKCNKKTTLTDLRYKVSGL 15222  
Qy 157 ----- 156  
Db 15223 TEGHEYEFIRMAENAGISAPSPSPFYKACDTVPKPGPGNPRVLDTSRSSISIANKNP 15282  
Qy 157 ----- 156  
Db 15283 IYDGSSEITGYMVEIALPEDEWQIVTPPAGLKATSYITITLTENQEKYKIRIYAMNSEGL 15342  
Qy 157 ----- 156  
Db 15343 GEPALVPGTPKAEDRMPLPPEIELDADLRKVVTIRACCTLRLFPVPIKGRPAPEVKWARDHG 15402  
Qy 157 ----- 156  
Db 15403 ESLDKASIESSTSSYTLIIVGNVNRFDGKYILITVENSNGSKSAFVNVVRVLDTPGPPODLK 15462  
Qy 157 ----- 156  
Db 15463 VKEVTKTSVTLTWDPPLLDGSGKIKNYIVEKRESTRKAYSTVATNCHKTSMKVDLOEGC 15522  
Qy 157 ----- 156  
Db 15523 SYFVRVLAENEYIGLPAETAESVKASERPLPGKITLMDVTRNSVLSWEKKEPHDGGSR 15582  
Qy 157 ----- 156  
Db 15583 ILGYIVEMQTKSGKWATCATVKVTEATITGLIOGEEYSFRVSAQNEKIGSDPQLSPVP 15642  
Qy 157 ----- 156  
Db 15643 IAKDLVTPPAFKLLFNFTVVLAGEDLKVDVFPFICRPTPAVTHKDNVPLKQTTTRVNAEST 15702  
Qy 157 ----- 156  
Db 15703 ENNSLLTIKDACREDVGHYVVVVKLTNSAGEALETNLVLDKPGPTGPKVKNDEVTADESIT 15762

QY	157	-----	156
Db	15763	LSWGPPIYDGGSSINNYIVEKRDTSITTWQIVSATVARTTIKACRLTKCEYQFRIAEN	15822
QY	157	-----	156
Db	15823	RYGKSTYLNSEPTVAQYPFKVPFGPTPVYTLSSRDSMEVQWNEPISDGGSRVIGYHLER	15882
QY	157	-----	156
Db	15883	KERNLSILWKLNKPTIPQTKFKTKTGLEGEVEYFRVSAENIVIGIKPSKVSECYVARDPC	15942
QY	157	-----	156
Db	15943	DPPGRPEALIVTRNSVTLQWKKPTYDGGSKITGYIVEKKELPEGRWMKASFTNIIDTHFE	16002
QY	157	-----	156
Db	16003	VTGLVEDHRYEFRVIAARNAAGVSEFSESTGAITARDEYDPPRISMDPKYKDTIVVHAGE	16062
QY	157	-----	156
Db	16063	SPKYVDADIYCKPIPTIQWIKGQDLSNTARLEIKSTDFATSLSVKDAVRVDSGNYILKAK	16122
QY	157	-----	156
Db	16123	NVAGERSVTNVKVLDRPPEGPVVISGVTAECTLANKPPLQDGGSDIINVIVERRET	16182
QY	157	-----	156
Db	16183	SRLWVTVDANVQTLSCVKTKLEGNEYTRIMAVNKYGVGEFLESEPVVAKNPVVPDA	16242
QY	157	-----	156
Db	16243	PKAPEVTITKDSMIVVWERPASDGGSEILGYVLEKRDKEGINWTRCHRLIGELRLVT	16302
QY	157	-----	156
Db	16303	GLIENHDEYFRVSAENAGLSEPPSAQKACDPIYKPGPNPKVIDITRSSVFLWS	16362
QY	157	-----	156
Db	16363	KPIYDGGCEIOGYIVEKCDVSGEWTMCPTPTGINKNTINEVEKLEKHEYNFRICAINKA	16422
QY	157	-----	156
Db	16423	GYGEHADVPPIIVEKLEAPDIDLDELRLKIINIRAGGSLRLFVPIKGRPTPEYKWKGV	16482
QY	157	-----	156
Db	16483	DGEIRDAALIDVTSSFTSLVDNVNRYDSGKYTLTLENSSGTSKSAFVTVRVLDTPSPVYN	16542
QY	157	-----	156
Db	16543	LKVTETKDSVITWBPPLLDGSKIKNYIVEKREATRSYAAVTVNCHNSWKIDQLOE	16602
QY	157	-----	156
Db	16603	GCSYFRVTAENEYIGLPAQTADPIKVAEVPQPGKITVDVTRNSVLSWTKPEHDGG	16662
QY	157	-----	156
Db	16663	SKIIQIVEMQAHSKSEKARVKSLOAVITNLTOGEEYLFVRVAVNEKGRSDPRSLAV	16722
QY	157	-----	156
Db	16723	PIVAKDLVIEPDVKPAFSSYSVQGDLEKIEVPISGRPKPTITWTKDGLPLKQTTIRYNV	16782
QY	157	-----	156
Db	16783	DSLDLTSLIKETHKDDGGQYGITVANVVGQKTASIEIIVTLQKPPPKPVKPFDDVSAES	16842
QY	157	-----	156

Db	16843	ITLSWNPPLYTGSCQITNVIVOKRDTTTTWDVVVSATVARTTLKVTKLKTGTEYQFRIFA	16902
QY	157	-----	156
Db	16903	ENRYGOSFALESPIVAQYPPYKEPGPTPFATAISKDSMVIQWHEPNNGGSPVIGYHL	16962
QY	157	-----	156
Db	16963	ERKERNLSILWTKVNTIIHDTQFKAQNLBEGIEYFRVVAENIVGVKASKNSECYVARD	17022
QY	157	-----	156
Db	17023	PCDPPCTPEPIMVKRNEITLQWTKPVYDGGSMITGVIVEKRDLPDGRWMKASFTNVIETQ	17082
QY	157	-----	156
Db	17083	FTVSGLTEDQRYEFRVIAKNAAGAIKSPSDSTGPITAKDEVELPRISMDFKFRDTIVVNA	17142
QY	157	-----	156
Db	17143	GETFRLEADYHGKPLPTIEWLRGDKIEBSARCEIKNTDFKALLIVKDAIRIDGGQYILR	17202
QY	157	-----	156
Db	17203	ASNAGSKGFPVNVKVLDRPPEGPVQVTGVTSEKSLTWSPLQDGGSDISHYVVVEKR	17262
QY	157	-----	156
Db	17263	ETSRLAWTVVASEVVTNSLKVTKLEGNEYVFRIMAVNKYGVGEPLASAPVLMKNPFVLP	17322
QY	157	-----	156
Db	17323	GPPKSLVETNIAKDSMTVCWNRPDSDGGSEIIGYIVEKRDERSIRWIKCNKRITDLRLR	17382
QY	157	-----	156
Db	17383	VTGLTEDHEVEYFRVSAENAGVGEPSPATVYKACDPVKFPGPTNAHIVDTTKNSITLA	17442
QY	157	-----	156
Db	17443	WCKPIYDGGSEILGYVVEICKADEEEMQIVTQOTGLRVTRFEISKUTEHQEYKIRVCALN	17502
QY	157	-----	156
Db	17503	KVCLGEATSVPGTVKPEDKLEAPELDLDSFLRKGIVVRAGGSARIHIPFKGRPTPEITWS	17562
QY	157	-----	156
Db	17563	REEGEFTDKVQIEKGVNYTQLSIDNCDNRNDAGKYILKLENSGSKSAFVTVKVLDTPGPP	17622
QY	157	-----	156
Db	17623	ONLAVKEVRKDSAFVWNEPPIIDGGAKVKNYVIDKRESTRKAYANVSSCKSKTSFKVENL	17682
QY	157	-----	156
Db	17683	TEGAIYYFRVMAENEFVGVPVETVDVAKAAEPPSPGKVTLTDSQTSASLAWKEPHEHD	17742
QY	157	-----	156
Db	17743	GGSRVLGYVVEMQPKTEKWSIVAESKVCNAVVTGLSSGOEYQFRVKAYNEKGS DPRVL	17802
QY	157	-----	156
Db	17803	GVPTIAKDLTIQPSLKLPTNTYSIQAGEDLKIEIPVIGRPRPNISWVKGEPLKQTRVN	17862
QY	157	-----	156
Db	17863	VEETATSTVLHIKEGNKDDFGKYTVTATNSAGTATENLSVIVLEKPGPPVGVPRFDEVA	17922
QY	157	-----	156

Db 17923 DFVWISWEPAYTGCQISNYIVEKRDTTTTTHHVMVSATVARTTIKITKLKTGTETQPRI 17982  
QY 157 ----- 156  
Db 17983 FAENRYGKSAPLDSKAVIVQYFKEPGPGPFVTSISKDQMLVQWHEPVNDGGTKIIGY 18042  
QY 157 ----- 156  
Db 18043 HLEQEKNSILWVLKLNKTPIDQTKFTTGLDEGLEVEFKVSAENIVGIGKPSKVCFA 18102  
QY 157 ----- 156  
Db 18103 RDCPPGRPEAIVITRNNTLKWKKPAYDGGKITGYIVEKKDLPDGRWMKASFTNVLE 18162  
QY 157 ----- 156  
Db 18163 TEFTVSLVEDQRYEFVRIARNAAGNFSEPSDSSCAITARDEIDAPNASLOPKYKDVIV 18222  
QY 157 ----- 156  
Db 18223 HAGETVLEADIRKPIPDVVWSKDGKELETAARMEIKSTIQKTTLVVKDCIRTDGGY 18282  
QY 157 ----- 156  
Db 18283 ILKLSNVGKTSIPITVKVLDROPPEGLKVTGVTAEKCYLAWNPPLQDGGANISHYII 18342  
QY 157 ----- 156  
Db 18343 EKRETSRLSWQVSTEVOALNYKVKLLPGNEYIFRVMVKNKYGIGEPLESGPVACNPY 18402  
QY 157 ----- 156  
Db 18403 KPPGPPSTPEVAITKDSMVVTWARPVDDGTEIEGYILEKRDKEGVWTKCNKKTLDL 18462  
QY 157 -----KGQPREPOVY----- 166  
Db 18463 RLVRTGLTEGHSYEFRAAENAGVGESEPSVYRACDALYPPGPPSNPKVTDTSRSV 18522  
QY 167 ----- 166  
Db 18523 SLANSKPIYDGGAPVKGYVEVEKEAADEWTTCTPPTGLOKQFTVTKLKENTEYNFRIC 18582  
QY 167 -----TLP----- 169  
Db 18583 AINSEGVGEPATLPGSVVAQERIEPPEIELDADLRKVVVLASATLRLFTVTKGRPEPEV 18642  
QY 170 ----- 169  
Db 18643 KWKAEGILTDRAQIEVTSFTMLVIDNVTRFDSGRYNLTLENNSGSKTAEFVNVRLDSP 18702  
QY 170 ----- 169  
Db 18703 SAPVNLITIREVKDSVILSWEPLLDGGAKITNIVEKRETRKAYATITNNCTKTTRI 18762  
QY 170 ----- 169  
Db 18763 ENLQEGCSYFRVLASNEYIGLPAETPEPVKVSEPPPLPPGRVTLVDVTRNTATIKWEK 18822  
QY 170 ----- 169  
Db 18823 ESDGSKITGYVEMQTKGSEKSTCTQVKTLEATISGLTAGEEYVFRVAAVNEKGRSDP 18882  
QY 170 ----- 169  
Db 18883 RQLGVPIARDIEIKPSVELPFHTFNVKAREQLKIDVPFKRQONTVNRKDGOTLKETT 18942  
QY 170 ----- 169  
Db 18943 RVNVSSKVTVTSLSIKEASKEDVGTYELCVSNSAGSITVPTITIVLDRPGPGPIRIDEV 19002  
QY 170 ----- 169  
Db 19003 SCDSITISWNPPEYDGGCQISNYIVEKETTSTTHHIVSQAVARTSIKIVRLTTGSEYQF 19062

QY 170 ----- 169  
Db 19063 RVCAENRYCKSSYSESSAVVAEYPPSPGPCTPKVHATKSTMLVTHQVPVNDGSRVI 19122  
QY 170 ----- 169  
Db 19123 GYHLEYKERSILMSKANKILIADTQMKVSGLDEGLMEYRYVYAENIAGIGKSKSCPVP 19182  
QY 170 PSRD----- 173  
|:|:  
Db 19183 PARDFCDPPGOPEVTNITRKSVSLKWSKPHYDGGAKITGYIVERRELDPGRWLKCNCTNI 19242  
QY 174 ----- 173  
Db 19243 QETYFEVTELTEDQRYEFVRFARNADSVSEPSSESTGIIIVKDDVEPPRVMDVKFRDVI 19302  
QY 174 ----- 173  
Db 19303 VVKAGEVLKINADIAGRPLPVISWAKDGEIEERARTEIISTDNHTLLTVKDCIRRDGTQ 19362  
QY 174 ----- 173  
Db 19363 YVLTILKNVAGTRSAVNCKVLDKPPAGPLEINGLTAEKCSLSWGRPQEDGADIDYI 19422  
QY 174 ----- 173  
Db 19423 VEKRETSHLAWTICEGELQMTSCKVTKLLKCNFYIFRVTGVNKGVGGEPLSVAILKALDP 19482  
QY 174 ----- 173  
Db 19483 FTVPSPTSLTITSVTKESMTLCWSRPESDGGSEISGYIIERREKNSLRWVRNKKPVYD 19542  
QY 174 ----- 173  
Db 19543 LRVKSTGLREGCEYERYVYAENAAAGLSLSETSLIRAEADVFLPSPPSKPKIVDSGKTT 19602  
QY 174 ----- 173  
Db 19603 ITIAWVPLFDGGAPITGYTVEYKKSDDTDMKTSIQSLRGTEYITISGLTTGAEYVFRVKS 19662  
QY 174 ----- 173  
Db 19663 VNVGASDPDSQDQIAKEREPEPLFDIDSEMRKTLIVKAGASFTMTVPERGPVNVL 19722  
QY 174 ----- 173  
Db 19723 WSKPDTDLRTRAYVDTTDSRTSLTIENANRNDGSKYTLTIQNVLSAASLTLLVVKVLDTPG 19782  
QY 174 ----- 173  
Db 19783 PPTNITVQDVTKEAVLSWDVPENDGGAPVKNYHIEKREASKAWSVTNNCNRLSYKVT 19842  
QY 174 ----- 173  
Db 19843 NLOEGAIYFRVSGENEFVGIPAETKEGVKITEKPSPEKLGVTISIKDSVSLTWLKE 19902  
QY 174 ----- 173  
Db 19903 HDGSRIVHYVVEALEKGQKNWVKCAVAKSTHHVVVSLGLENSEYFFRVFAENQAGLSOPR 19962  
QY 174 ----- 173  
Db 19963 ELLLPVLKEOLEPPEIDMKNFPSHTVVVRAGSNLKVDIPISGKPLPKVTLSDRGVPLKA 20022  
QY 174 ----- 173  
Db 20023 TMRFNTETAENLTINLKESVTADAGRYEITAANSSGTTRAFINIVVLDREPPTGPVVI 20082  
QY 174 ----- 173  
Db 20083 SDITEESVTLKWEPPKYDGGSQVTNYILLKRETSATVTEVSATVARTMMKMLTTGEE 20142

QY 174 ----- 173  
Db 20143 YQFRIKAENRFGISDHIDSACVTKLPYTPPGPSTPWNTVNTRESITVGMHEPVSNNGS 20202  
QY 174 ----- 173  
Db 20203 AVVGYHLEMKDRNSILWQANKLVIRTHPKVTTISAGLIYEFVRYAENAAAGVKPSHPS 20262  
QY 174 ----- 182  
          -ELTKNOVSL-  
          :::| | | |  
Db 20263 EPVLAIDACEPPRNRVITDISKNSVLSWQOPAFDGGSKITGYIVERRDLPDGRWTKASF 20322  
QY 183 ----- 185  
          -TCL-  
          | | |  
Db 20323 TNVTEQFIISGLTONSOYEFVRFARNAVGSI SNPSEWGPITCIDSYGGPVIDLPLEYT 20382  
QY 186 ----- 185  
Db 20383 EVVKYRAGTSVKLRAGISGKPAFTIEWYKDDKELOTNALVCVENTTDLASILKADRLN 20442  
QY 186 ----- 185  
Db 20443 SGCYELKLRNAGSASATIRVQILDKPGPPGPIEFKVTVAEKITLLWRPPADGGAKIT 20502  
QY 186 ----- 185  
Db 20503 HYIVEKRETSRVVWSMVSEHLEECIITTKIKGNEYIFRRAVNKYGIGEPLESDSVA 20562  
QY 186 ----- 185  
Db 20563 KNAFVTPGPGPIPEVTKITKNSMTVYVMSRPIADGGDISGYFLEKDKKSLGWFKVLKET 20622  
QY 186 ----- 185  
Db 20623 IRDTRQKVTGLTENSIOYRVCANRAGQGFSEPSEFYKAADPIDPPGPPAKIRIADST 20682  
QY 186 ----- 185  
Db 20683 KSSITLGSKPVDGGSATGYVVEIRQEGBEWTTVSTKGVEVTEYVYVSNLKPGVNY 20742  
QY 186 ----- 185  
Db 20743 FRVSANACOGGPIEMNEPVQAKDILEAPEIDLVALRTSVIAKAGEDVQVLIPEKGRP 20802  
QY 186 ----- 185  
Db 20803 PPVTVWRKDEKNLGS DARYSIENTDSSLLTIPQVTRNDTGKYLITENGVGEPKSTVS 20862  
QY 186 ----- 185  
Db 20863 VKVLDTPAACQKLQVHVSRGVTLLWDPLIDGGSPINNVIEKRDATKRTWSVYSHK 20922  
QY 186 ----- 185  
Db 20923 SSTSKLIDLSEKTPFRVLAENEIGEPCEPTEPVKAAEVPAPIRDLMSKDKSTKTSV 20982  
QY 186 ----- 185  
Db 20983 ILSWTRKDPDGGSVITEYVVVERKKGQETWSHAGISKTCIEIVSQLKEQSVLEFRFAKN 21042  
QY 186 ----- 185  
Db 21043 EKLSDPVTIGITVTKELIITPEVDLSDIPGAQVTVRIGHNVHLELPYKPKPSISWLK 21102  
QY 186 ----- 185  
Db 21103 DGLPKSEFVRSFKTENKITLSIKNAKHBGGKYTVILDNAVCRVIAPIVITVTLGPPSK 21162  
QY 186 ----- 185  
Db 21163 PKGPIRDEIKADSVILSDVPEDNGGGETCYTSIEKRETSQTNWKMVCSSVARTPKVP 21222  
QY 186 ----- 185

Db 21223 NLVKDAEYQFRVRAENRYGVSQLVSSIIIVAKHQFRIPGPKPVYINVTSDGMSLTWDA 21282  
QY 186 ----- 190  
          -VKGFY-  
          | | |  
Db 21283 PVDGGSVETGPHVEKKEKERNILWQKNTSPISGREYRATGLVEGLDYQFRVYAENSAGL 21342  
QY 191 ----- 193  
          --PSD-  
          | | |  
Db 21343 SSFSDPSKFTLAVSPVDPGTPDYIDVTRITILKNPPLDGGSKIVGYSIEKROGNER 21402  
QY 194 ----- 193  
Db 21403 WVRCNFTDSECOYTVTGLSPGDRYEFRIIARNNAVGTISPPSQSGIIMTRDENVPPIVE 21462  
QY 194 ----- 193  
Db 21463 FGPEYFDGLIIKSGESURIKALVOGRVPVPTWFKDGVETEKRMMEITDVLGSTSIFVR 21522  
QY 194 ----- 193  
Db 21523 DATRDHRGVTVVEAKNAGSAAEIKVKVQDTPGKVVGPITFTNITGKMTLWWDAPLND 21582  
QY 194 ----- 193  
Db 21583 GCAPITHYIIEKRETSRLAWALIEDKCEAQSYTAIKLINGNEYQFRVSANVKFGVRPLD 21642  
QY 194 ----- 193  
Db 21643 SDPVVAIOIYTVDPADGPIEPSNITGNSITLTWARPESDGGSETOQYILERREKKSTRWV 21702  
QY 194 ----- 193  
Db 21703 KVISKRPISETRKFVKTGLTEGNEYEFHVMAENAGVGPASGISRLIKREPVPNPPGPTV 21762  
QY 194 ----- 198  
          -IAVEW-  
          :::| | |  
Db 21763 VKVTDTSKTTVSLWSKPVDFDGGMEIIGIEMCKADLGDWHKVNAEACVKTRYVTDLQ 21822  
QY 199 ----- 198  
Db 21823 AGEYKFRVSAINGAGKGDSCVETGTIKAVDRLTAPELDIDANFKQTHVVRAGASIRLFI 21882  
QY 199 ----- 201  
          -ESN-  
          | | |  
Db 21883 AYQGRPTTAVWSKPSDNLRLADIHTDSEFTLTVCNCRNDACKYTLTVENNSGSKSI 21942  
QY 202 ----- 201  
Db 21943 TFTVKVLDTPGPPGPIITFKDVTGRSATLMDAPLLDGGARITHYVVEKREASRSQWVIS 22002  
QY 202 ----- 201  
Db 22003 EKCTROIKVNDLAEGVYYFRVSANVEYGVGEPEYEMPEPIVATEQPAPRRLDVDTSK 22062  
QY 202 ----- 201  
Db 22063 SSAVLAWLKPDHDDGSRITGYLLEMRKOGSDFWVEAGHTKOLTFTVERLVEKTEYEFVK 22122  
QY 202 ----- 201  
Db 22123 AKNDAGYSEPREAFSSVIIKEPQIEPTADLTGINTQLITCKAGSPFTIDVPISGRPAKV 22182  
QY 202 ----- 201  
Db 22183 TWKLEEMRLKETDRVSITTTTKDRITTLAVKDSMRGDSGRYFLTLENTAGVKTFSVTVVIG 22242  
QY 202 ----- 204  
          --QOP-  
          | | |  
Db 22243 RGPVVTGPVIEVSSVSAESCVLSWGEKPKDGGGTEITNIVKRESGTTAWQLVNSVVKRTQ 22302  
QY 205 ----- 204

```
Db 22303 IKVTHLTMYEYFRVSSNRFVSKPLSAPIAIEHPFPSPAPTRPEVYHVSANMSI 22362
QY 205 -----ENNYKTT----- 211
Db 22363 RWEBPYHDGSKIIGYWEKKEKERNILWKNKVCLECNKYKVTGLVEGLEYOFRYALN 22422
QY 212 ----- 211
Db 22423 AAGVSKASRPIMAQNPDAPGRPEVTDVTRSTVSLWSAPAYDGGSKVGYIIERP 22482
QY 212 ----- 211
Db 22483 VSEVGDGRWLKCNKTYIVSDNFFVTALSEGDTYEFRLAKNAAGVISKSESTGPVTCRD 22542
QY 212 --PP----- 213
Db 22543 EYAPPKAELDARLHGLVLTIRAGSDLVLDAAVGGKPEPKIITWKGDKELDCEKVSLOVT 22602
QY 214 -----VLDSDG----- 219
Db 22603 GKRAVIAKFCDRSDSGKYTLTKNASGTKAVSVHVKVLDSPGCKLTVSRVTOEKCTL 22662
QY 220 ----- 219
Db 22663 AWSLPQEDGGAETHYIIVERRETSRLNMVIVEGECPTLSYVVTRLIKNNEYIFRVRVANK 22722
QY 220 -----SFFL----- 223
Db 22723 YCPGVPESEPIVARNSTIPSPGPPEVGTGKEHIIOWTKPESDGNESINVLVDKR 22782
QY 224 ----- 223
Db 22783 EKKSLRWTRVNKDYVYDTRLKVTSLMEGCDYQFRVTA^NAGNSEPSEASNFISCREPS 22842
QY 224 ----- 223
Db 22843 YTPGPPSAPRVVDTKHSLSLAWTKPMYDGGTDIVGYVLEMQEKTDQWYRVHTNATIRN 22902
QY 224 ----- 223
Db 22903 TEFTVPDLKMGQKYSFRAAVNVKMGSEYSESIAEIPVERIEIPDLELADDLKKTVTIR 22962
QY 224 -----YKLTVDK----- 231
Db 22963 AGASLRLMVSVGRPPPVITWSKOGIDLASRAIIDTTESYSLIIVDKVNRVDAGRYTIEA 23022
QY 232 ----- 231
Db 23023 ENQSGKKSATVLVKVYDTPGCPSPVKVEVSRDSVTITWEIPTIDGGAPVNNYIVEKREA 23082
QY 232 ----- 231
Db 23083 AMRAFKTVTCKSKTYLRISGLVEGTMYFVRLPENIYIGEPCEPSETSDAVLVSEVPLPA 23142
QY 232 ----- 231
Db 23143 KLEVVDVTKSTVLAWERPLYDGGSLRATGYVLEACKAGTERMKVYVTLKPTVLEHTVTSL 23202
QY 232 ----- 231
Db 23203 NEGEQYLFIRAQNEKGVSEPRETVTAVTVQDLRLVLPIDLSLTPMOKTIHVPAGRPVELV 23262
QY 232 ----- 231
Db 23263 IPIAGRPPPAASWFFAGSKLRESERVTVETHKVAKLTIRETITRDTEYTLKKNVTGT 23322
QY 232 ----- 231
Db 23323 TSETIKVILDKPGPTGPIKIDEIDATSITISWEPPELDGGAPLSGYVWEQDAHRPGW 23382
QY 232 ----- 231
Db 23383 LPVSESVTRSTFKFTRLLEGNEYVFRVAATNRFNGIGSYLQSEVIECRSSIRIPGPPETLQ 23442
```

```
QY 232 ----- 231
Db 23443 IPDVSBDGMLTWYPPEDDDGGSQVTGYIVERKEVRADRRVRVKNKVPVMTYRSTGLTEG 23502
QY 232 ----- 231
Db 23503 LEYEHRTVAINARGSGKPSRPSKPIVAMDPIAPPCKPQNPRVTDTRTTSVSLAWSVPDE 23562
QY 232 ----- 231
Db 23563 GGSKVTGYLIEQKVDQHEWTKCNTPTTKIREYTLTHLPQGAERYFRVLACNAGGPPEA 23622
QY 232 -----SRWOG----- 237
Db 23623 EYPGTVKVTMELEYPDYELDERYQEGIFVRQGGVIRLTIPIKKGKPPICKWTKEGQDISK 23682
QY 238 ----- 237
Db 23683 RAMIATSETHTELVIKEADRGDSGTYDLVLENKCGKAVYIKVRVIGSPNSPEGPLEVDD 23742
QY 238 ----- 237
Db 23743 IOVRSVRVSWRPPADDDGADILGYILERREVPKAAWYITIDSKVRGTSLVVKGLENVEYH 23802
QY 238 ----- 237
Db 23803 FRVSAENQFGISKPLKSEPVTPKTLNPPPEPPSNPPEVLDVTKSSVLSNSRPKDDGGS 23862
QY 238 ----- 237
Db 23863 RVTGYIIERKETSTDKWRHNKQITTTMYTVTGLVPAEYQFRIIAONDVGLSETSPAS 23922
QY 238 ----- 237
Db 23923 EPVCKDPDPKSPQCELEILISKDSVTLQWEKPECDDGKRGILGYWVEYRQSGDSANWK 23982
QY 238 ----- 237
Db 23983 SNKERIKDKQFTIGGLEATEYEFVRFAENETGLSRPRRTAMSIKLTSGEAPGIRKEM 24042
QY 238 ----- 237
Db 24043 KDVTTKLGAAQLSCQIVGRPLPDIKYRFGKELIQSRKYMSSDGRHTLTVMTFEQED 24102
QY 238 ----- 237
Db 24103 EGYVTCIATNEVGEVETSSKLLLOATPOFHPGYPLKEXYGAAGVSTLRLHVMYIGRPVPA 24162
QY 238 -----NVF----- 240
Db 24163 MTFHGOQLQNSENITIENTEYTHLVKMNQVQKTHAGKYKQLSNVFGTVDAILDVEI 24222
QY 241 ----- 240
Db 24223 QDKPDKPTGPIVIEALLKNSAVISWKPADDGGSWITNYVVEKCEAEGAEHQLVSSAIS 24282
QY 241 ----- 240
Db 24283 VTTCRIVNLNENAGYFRVSAQNTFGISDPLEVSSVLIKSPFEKPGAPKPTITAVTKD 24342
QY 241 SCV----- 244
Db 24343 SCVVAWKPPASDGGAKIRNYLEKREKKONKWSVTTTEIRETVFSVKNLTIEGLEFEVR 24402
QY 245 ----- 244
Db 24403 KCENLGSESEWSEITSEPTPKSDVPIQAPHKFEELNLRVQSNATLVCKVTHGPKPIV 24462
QY 245 ----- 244
Db 24463 KWRQKKEIADGLKRYIQEFPKGGYHQLIIASVTDODDATVYQVRATNOGGSVSGTASLEV 24522
```



```
QY 245 ----- 244
Db 24523 EVPAKIHLPKTLGMAVHALGEVWSIKIPFSGKDPVITWQGDLDINNGHYQVIVT 24582
QY 245 ----- 244
Db 24583 RSFTSLVFPNGVERKDAGFYVVCANKNRFIDQKTVELDVADVPDPGRGVKVSVDVSRDSVN 24642
QY 245 ----- 244
Db 24643 LWTWPASDCGSKITNYIIVEKCATTAERWLRVQARETRYVINLFGKTSYQFRVIAENK 24702
QY 245 ----- 244
Db 24703 FGLSKPSEPSEPTITKEDKTRAMNYDEVDREVSMTKASHSTKELYEYMIADLGR 24762
QY 245 ----- 248
Db 24763 GEFIVHRCVETSSKTYMAKFVKVKGTDQVLVKEISILNIARHRNHLHSHESFESMEE 24822
QY 249 ----- 252
Db 24823 LVMIFEISGLDIFERINTSAFELNBEREIVSVHVQVCEALQFLHSHNIGHFDIRPENIY 24882
QY 253 ----- 252
Db 24883 QTRRSSTIKIIEFGARQLKPGDNFRLLFTAPEYAPEVHQHDVSTATDMWSLGLTVYV 24942
QY 253 ----- 254
Db 24943 LLSGINPFLAETNOQIENIMNAEYTFDEAFKEISTEAMDFVDRLLVKKRKSMTASEA 25002
QY 255 ----- 261
Db 25003 LQHPWLKQKIERVSTKVIWRLKHRRYHTLKKDLNMYVSAARISCGGAIKRSQGVSAK 25062
QY 262 ----- 261
Db 25063 KVASIEIGPVGQIMHAVEGEGHVKYCKIENYDQSTQVWYFGVRLSENSEKEITY 25122
QY 262 ----- 261
Db 25123 EDGAILYVKDITKLDGTYRCKVWNDYGEDSSYAELFKGVREVYDYCRRTMKKIKRR 25182
QY 262 ----- 261
Db 25183 TDTMRLLERPPEFTPLYNKTAIVGENVRGVITIVHPEPHVTWYKSGQIKPGDNDKKY 25242
QY 262 ----- 261
Db 25243 TFESDKGLYQLTINSVTDDAEVTVVARNKYGEDSCKAKLTVTLHPPPTDSTLRPMFKR 25302
QY 262 ----- 261
Db 25303 LLANAEQEGQSVCFEIRVSGIPPPPTLKWEKGQPLSLGNIEIIEGLDYALHIRTDL 25362
QY 262 ----- 261
Db 25363 PEDTGYRVATNTAGSTSCQAHQVRLRYKKQEPKSEEHVHQKQIDKTLRMAEIL 25422
QY 262 ----- 261
Db 25423 SGTESVPLTQVAKALREAAVLYKPAVSTKTVKGEFRLEIEEKKERKLMPYDVPEPRK 25482
QY 262 ----- 264
Db 25483 YKOTTIEEDORIKQFVPMDSMKWYKKIRDQYEMPGK 25518
```



CC form heterodimers (SR-alpha:beta-1) that act as antagonist to the  
CC cytokine by binding the cytokine to form a non-functional complex. The  
CC receptor components are shared by cytokines such as the CNTF (ciliary  
CC neurotrophic factor) family of cytokines. The invention provides the  
CC basis for the development of IL-6 antagonists, as they show that if, in  
CC the presence of a ligand, a non-functional intermediate complex,  
CC consisting of the ligand, its alpha receptor and its beta-1 receptor  
CC component, can be formed, it will effectively block the action of the  
CC ligand. Effective antagonists of IL-6 or CNTF consist of heterodimers  
CC of the extracellular domains of the alpha specificity determining  
CC components of their receptors and the extracellular domain of gp130.  
CC The resultant heterodimers, function as high-affinity traps, rendering  
CC the cytokine inaccessible to form a signal transducing complex with the  
CC native membrane-bound forms of their receptor. The nucleic acids and  
CC polypeptides are useful for treating cytokine-related diseases or  
CC disorders such as osteoporosis and primary and secondary effects of  
CC cancer including multiple myeloma or cachexia.  
XX  
SQ Sequence 690 AA;  
Query Match 94.0%; Score 1254; DB 21; Length 690;  
Best Local Similarity 52.1%; Pred. No. 6.1e-36;  
Matches 233; Conservative 4; Mismatches 8; Indels 202; Gaps 1;  
QY 2 WLKAFY----- 7  
DB 244 WNSFYRLRFLRYRAERSKFTTWMVKDLQHHCVHDAWSGLRHVVQLRAQEEFGQGW 303  
QY 8 ----- 7  
DB 304 SEWSPAMGTPWESRPPAENEVSTPMQALTNKDDNILFRDSANATSLPVQAGAST 363  
QY 8 ----- 7  
DB 364 KGPSVFPLAPSSKSTSGGTAALCLVKDYFPEPTVSWNSGALTSGVHTFPAVLQSSGLY 423  
QY 8 -----DKVAEKLKEAFMDKTHRCPCPAPPELLGGPSV 39  
DB 424 SLSSVVTVPSSSLGTQYTCNVNHRKPSNTKDKKVEPKSCDKTHRCPCPAPPELLGGPSV 483  
QY 40 FLFPKPKDTLMISRTPEVTCVVDVSHEDPEVKFNWYVDGVEVHNATKPREQYNSTY 99  
DB 484 FLFPKPKDTLMISRTPEVTCVVDVSHEDPEVKFNWYVDGVEVHNATKPREQYNSTY 543  
QY 100 RVVSVTLVLHQDLNGKEYCKVSNKALPAPIEKTISKAKGQPREPQVYTLPPSRDELTK 159  
DB 544 RVVSVTLVLHQDLNGKEYCKVSNKALPAPIEKTISKAKGQPREPQVYTLPPSRDELTK 603  
QY 160 NOVSLTCLVKGFYPSDIAVEWESNGOPENNYKTTTPVLDSDGSFFLYSKLTVDKSRWQQG 219  
DB 604 NOVSLTCLVKGFYPSDIAVEWESNGOPENNYKTTTPVLDSDGSFFLYSKLTVDKSRWQQG 663  
QY 220 NVFSCVMHEALHNHYTQKSLSLSPCK 246  
DB 664 NVFSCVMHEALHNHYTQKSLSLSPCK 690  
RESULT 13  
AAW70796  
ID AAW70796 standard; protein; 859 AA.  
XX  
AC AAW70796;  
XX  
DT 03-FEB-1999 (first entry)  
XX  
DE Human gp130-Fc-His6 amino acid sequence.  
XX  
KW gp130; cytokine antagonist; interleukin; gamma-interferon;  
KW granulocyte macrophage colony-stimulating factor; J peptide;  
KW transforming growth factor-beta.  
XX  
OS Synthetic.  
OS Homo sapiens.

XX Key Location/Qualifiers  
FH Protein 1..619  
FT /note= "human gp130"  
FT Misc-difference 2  
FT /label= L2V  
FT /note= "amino acid changed to accommodate a Kozak  
FT sequence"  
FT Peptide 1..22  
FT /note= "signal peptide"  
FT Misc-difference 620..621  
FT /note= "Ser-Gly bridge"  
FT Protein 662..853  
FT /note= "from the Fc domain of human IgG1"  
FT Disulfide-bond 632..635  
FT Peptide 854..859  
FT /note= "hexahistidine tag"  
XX  
PN US5844099-A.  
XX  
PD 01-DEC-1998.  
XX  
PF 27-NOV-1995; 95US-0563105.  
XX  
PR 27-NOV-1995; 95US-0563105.  
PR 20-OCT-1993; 93US-0140222.  
XX  
PA (REGE-) REGENERON PHARM INC.  
XX  
PI Economides A, Stahl N, Yancopoulos GD;  
XX  
DR WPI; 1999-044669/04.  
XX  
PT Cytokine antagonists - comprising extracellular domains of  
PT specificity-determining and signal-transducing components of  
PT cytokine receptor  
XX  
PS Example 3; Fig 4; 46pp; English.  
XX  
CC The present sequence represents the amino acid sequence of human  
CC gp130-Fc-His6. The protein is used in the course of the invention. The  
CC specification describes cytokine antagonists comprising only the  
CC extracellular domain of the specificity-determining component of  
CC the cytokine receptor and the extracellular domain of a  
CC signal-transducing component of the cytokine receptor. The cytokine  
CC is an interleukin (IL-1, IL-2, IL-3, IL-4, IL-5 or IL-15),  
CC granulocyte macrophage colony-stimulating factor (GM-CSF),  
CC gamma-interferon or transforming growth factor-beta (TGF-beta). The  
CC antagonist is capable of binding the cytokine to form a nonfunctional  
CC complex. The compounds have therapeutic activity as cytokine antagonists  
CC and can also be used in assays for identifying novel agonists and  
CC antagonists of cytokines.  
SQ Sequence 859 AA;  
Query Match 94.0%; Score 1254; DB 20; Length 859;  
Best Local Similarity 60.5%; Pred. No. 1e-35;  
Matches 233; Conservative 8; Mismatches 5; Indels 139; Gaps 3;  
QY 1 DW-----LKA----- 6  
DB 469 DWQEDGTVHRTYLRGNLAESKCYLITVPYADGPGSPESIRAYLQAPPKGPVTRK 528  
QY 7 -----YDKVAEKLKEAF----- 19  
DB 529 KVGKNEAVLEWDLQDPVDVQNGFIRNYTIFRTIIGNETAENVNDSSTHTYTLSSLTSDTL 588  
QY 20 -----DKTHCTPPCPAPPELLGGPSVFL 41  
DB 589 MYRMAAYTDEGGKGDPFTTTPKFAQGEIESGEPKSCDKTHCTPPCPAPPELLGGPSVFL 648  
QY 42 FPPKPKDTLMISRTPEVTCVVDVSHEDPEVKFNWYVDGVEVHNATKPREQYNSTYRV 101  
OS  
OS

XX AAW48650;  
XX 04-AUG-1998 (first entry)  
DT Heavy chain of hmb425 fused to TNF alpha.  
DE Antibody-cytokine fusion protein; tricistronic vector; chimeric;  
KW TNF alpha; IL-2; IRES; internal ribosome entry site.  
XX Synthetic.  
OS Homo sapiens.  
XX  
XX Key Location/Qualifiers  
FH Region 1..494  
FT /note= "Heavy chain of human mAb 425"  
FT Region 495..652  
FT /note= "TNF alpha"  
XX  
PN WO9811241-A1.  
XX  
XX 19-MAR-1998.  
XX  
XX 02-SEP-1997; 97WO-EP04765.  
XX  
XX 30-SEP-1996; 96EP-0115635.  
PR 16-SEP-1996; 96EP-0114820.  
XX  
XX (MERE ) MERCK PATENT GMBH.  
XX  
XX Bruemmer W, Burge C, Dunker R, Hauser H, Mielke C;  
PI Rieke E, Von Hoegen I, Welge T;  
XX  
XX WPI; 1998-207400/18.  
DR N-PSDB; AAV18096.  
XX  
XX Oligo:cistronic expression vector - useful for production of, e.g.  
PT MAb425/TNF- $\alpha$  or MAb425/IL-2 antibody fusion protein  
XX  
XX Disclosure; Fig 15; 89pp; English.  
XX  
XX The present sequence represents a fusion protein comprising of TNF  
CC alpha fused to the C-terminus of the heavy chain of the human  
CC monoclonal antibody 425 (hmb425). The hmb425 has specificity for  
CC the human EGF receptor. The invention claims for a new pMCDHAP  
CC tricistronic vector (AAV18096) for the expression of an  
CC antibody-cytokine fusion protein, hmb425-TNF alpha. The TNF alpha  
CC sequence can be substituted by the IL-2 sequence. The vector also  
CC contains a strong promoter/enhancer unit, a selection marker gene and at  
CC least two poliovirus derived internal ribosomal entry site (IRES)  
CC sequences. The vector can be expressed in mammalian host cells for the  
CC production of heteromeric fusion proteins. This expression system is  
CC claimed to produce the heteromeric proteins in high yields.  
XX  
SQ Sequence 652 AA;  
Query Match 94.0%; Score 1254; DB 19; Length 652;  
Best Local Similarity 54.4%; Pred. No. 5.3e-36;  
Matches 234; Conservative 5; Mismatches 7; Indels 184; Gaps 2;  
QY 1 DWLKF-----YD- 8  
DB 65 EWIGFNPSNGRTNYNEKFSKATMTVDSTNTAYMELSLRSEDVAVYCASRDYDYG 124  
QY 9 ----- 8  
DB 125 RYFDYWGQGLTVSSGEWILCAWALCPTPRSHGTTSLAASTKGPSVFLAPSSKSTSG 184  
QY 9 ----- 8  
DB 185 GTAAIGCLVKDYFPEPVTWNSNGALTSGVHTFPAVLQSSGLYSLSVVTVFSSSLGTGT 244  
QY 9 -----KVAEKLKAEAFMDKTHTCPPCPAPPELLGGPSVFLFPPKPKDTLMISRTTP 56

DB 245 YICNVNHPSTNTKVDKKVEPKSCDKTHTCPPCPAPPELLGGPSVFLFPPKPKDTLMISRTTP 304  
QY 57 EYTCVVVDVSHEDPEVKFNWYVDGVEVHNAKTKPREEQYNSTYRVVSVLTVLHQDWLNGK 116  
DB 305 EYTCVVVDVSHEDPEVKFNWYVDGVEVHNAKTKPREEQYNSTYRVVSVLTVLHQDWLNGK 364  
QY 117 EYCKYVKSNKALPAPIEKTISKAKGQPREPOVYTTLPSSRDELTKNOVSLTCLVKGFYPSDI 176  
DB 365 EYCKYVKSNKALPAPIEKTISKAKGQPREPOVYTTLPSSRDELTKNOVSLTCLVKGFYPSDI 424  
QY 177 AVESWGQPNENNYKTPPVLDSDGSGFFLYSKLTVDKSRWQQGNVSCSYMHHEALHNHYT 236  
DB 425 AVESWGQPNENNYKTPPVLDSDGSGFFLYSKLTVDKSRWQQGNVSCSYMHHEALHNHYT 484  
QY 237 QKSLSLSPGK 246  
DB 485 QKSLSLSPGK 494  
RESULT 12  
AA92195  
ID AAY92195 standard; protein; 690 AA.  
XX  
XX AC AAY92195;  
XX  
DT 01-AUG-2000 (first entry)  
XX  
XX Human IL-6R-alpha-C-gamma-1 fusion protein.  
DE  
KW IL-6R-alpha-C-gamma-1; cytokine; antagonist; CNTF; receptor;  
KW fusion protein; cytostatic; immunomodulator; osteopathic.  
XX  
OS Synthetic.  
OS Homo sapiens.  
XX  
FH Key Location/Qualifiers  
FT Protein 1..358  
FT /label= IL-6R-alpha  
FT Peptide 359..360  
FT /note= "Ala-Gly bridge"  
FT Protein 361..690  
FT /label= C-gamma-4  
XX  
PN WO200018932-A2.  
XX  
PD 06-APR-2000.  
XX  
XX 22-SEP-1999; 99WO-US22045.  
PR 25-SEP-1998; 98US-0101858.  
PR 19-MAY-1999; 99US-0313942.  
XX  
XX (REGE-) REGENERON PHARM INC.  
XX  
XX Stahl N, Yancopoulos GD;  
XX WPI; 2000-293165/25.  
XX  
PT Isolated nucleic acid molecule for treating cytokine-related diseases  
PT or disorders encodes a fusion polypeptide capable of binding a cytokine  
PT to form a nonfunctional complex  
XX  
PS Example 4; Page -: 152pp; English.  
XX  
CC The invention concerns production of antagonists to any cytokine that  
CC utilizes an alpha specificity determining component, which when combined  
CC with the cytokine, binds to a first beta signal transducing component to  
CC form a non-functional intermediate which then binds to a second beta  
CC signal transducing component causing beta-receptor dimerization, the  
CC soluble alpha specificity determining component of the receptor  
CC (sR-alpha) and the extracellular domain of the first beta signal  
CC transducing component of the cytokine receptor (beta-1) are combined to

```
PN WO200123431-A1.
XX 05-APR-2001.
XX 29-SEP-2000; 2000WO-JP06775.
XX 30-SEP-1999; 99JP-0278292.
XX (KYOW ) KYOWA HAKKO KOGYO KK.
XX Hanai N, Nakamura K, Niwa R;
XX WPI; 2001-266142/27.
XX Monoclonal antibodies against ganglioside GM2 combined with drugs,
XX radioisotopes or proteins for treatment and diagnosis of cancer -
XX Claim 43; Page 61-65; 80pp; Japanese.
XX The present invention relates to derivatives of an antibody against
XX ganglioside GM2. The antibody may be a monoclonal antibody or its
XX fragments. The antibody is combined with a radioactive isotope,
XX protein or small drug in the treatment and diagnosis of cancer.
XX Sequence 583 AA;
XX
Query Match 94.1%; Score 1255; DB 22; Length 583;
Best Local Similarity 56.2%; Pred. No. 3.8e-36;
Matches 234; Conservative 5; Mismatches 7; Indels 170; Gaps 2;
Qy 1 DWLK-----AF 6
Dy 35 DWVKQSQGGLEWGMGIYPNNGGTGYNOKFKSKVTITVDTSTAYMELHSLRSEDYAVY 94
Qy 7 Y-----7
Dy 95 YCATYGHYGYMFAYWGGTLTVSSASTKGPSVFPLAPSSKSTSGGTAALGCLVKDYFP 154
Qy 8 -----DKV 10
Dy 155 EPVTVMNSGALTSGVHTFPAVLQSSGLYSLSSVTVTPSSSLGTQTYICNVNHRKPSNTKV 214
Qy 11 AEKLEAFMDKTHCPCPAPPELLGGPSVFLFPPPKDTLMISRTPEVTCVVDVSHEDP 70
Dy 215 DKKVEPKSCDKHTHTCPCPAPPELLGGPSVFLFPPPKDTLMISRTPEVTCVVDVSHEDP 274
Qy 71 EVKFNWYVDGVEVHNAKTPREEQYNSTYRVVSVLTVLHQDWLNGKEYCKVKSNKALPAP 130
Dy 275 EVKFNWYVDGVEVHNAKTPREEQYNSTYRVVSVLTVLHQDWLNGKEYCKVKSNKALPAP 334
Qy 131 IEKTSKAGQPREPQVYTLPPSRDELTKNOVSLTCLVKGYPSDIAVWESNGOPENNY 190
Dy 335 IEKTSKAGQPREPQVYTLPPSRDELTKNOVSLTCLVKGYPSDIAVWESNGOPENNY 394
Qy 191 KTTPLVLDSDGSFFLYSKLTVDKSRWQOGNVFSCSVMEALHNNHYTKQSLSPCK 246
Dy 395 KTTPLVLDSDGSFFLYSKLTVDKSRWQOGNVFSCSVMEALHNNHYTKQSLSPCK 450
RESULT 10
AAG64473
ID AAG64473 standard; Protein: 473 AA.
XX
AC AAG64473;
XX
DT 25-SEP-2001 (first entry)
XX
DE Human type antihuman IgE antibody H chain 3.
XX
KW Human; antihuman IgE antibody; immunoglobulin; treating;
XX allergic disease.
XX Homo sapiens.
OS
```

```
XX Key Location/Qualifiers
FH Misc-difference 22
FT /note= "Encoded by CAG"
ET Misc-difference 129..141
FT /note= "Encoded by ccgtggggcc agggaaacacc ggtcgcccttt
XX gactacgtc"
XX WO200151507-A1.
XX 19-JUL-2001.
XX 15-JAN-2001; 2001WO-JP00181.
XX 14-JAN-2000; 2000JP-0007061.
XX (SNOW ) SNOW BRAND MILK PROD CO LTD.
XX Washida N, Takahashi K, Satake T, Fujise N, Tanaka H, Kuriyama M;
XX WPI; 2001-442132/47.
XX N-PSDB; AAH47901.
XX New peptide used for screening human anti-human immunoglobulin E
XX monoclonal antibody useful in medical compositions for treating
XX allergies -
XX Example 6; Page 59-60; 70pp; Japanese.
XX The present sequence is that of a human type antihuman IgE antibody H
XX chain. The invention relates to a peptide useful in a method for
XX screening for human type antihuman IgE monoclonal antibodies
XX (AAH47897-AAH47904 encoding AAG64469-AAG64476) useful for preventing
XX and/or treating allergic disease.
XX Sequence 473 AA;
XX
Query Match 94.0%; Score 1254; DB 22; Length 473;
Best Local Similarity 55.5%; Pred. No. 2.5e-36;
Matches 233; Conservative 6; Mismatches 7; Indels 174; Gaps 2;
Qy 1 DWLK-----4
Dy 54 NWROVPQGGLEWAGIIPFGPKYAKFGQGRVSLTADRSTNTAYMEMARLSRSDTAVY 113
Qy 5 -----AFYD-----8
Dy 114 YCARAPTLYDTEGTGPFYWGOGTPVAVSSASTKGPSVFPLAPSSKSTSGGTAALGCLVK 173
Qy 9 -----8
Dy 174 DYFPEPVTVMNSGALTSGVHTFPAVLQSSGLYSLSSVTVTPSSSLGTQTYICNVNHRPS 233
Qy 9 --KVAELKEAFMDKTHCPCPAPPELLGGPSVFLFPPPKDTLMISRTPEVTCVVDVS 66
Dy 234 NTKVDKVEPKSCDKHTHTCPCPAPPELLGGPSVFLFPPPKDTLMISRTPEVTCVVDVS 293
Qy 67 HEDPEVKFNWYVDGVEVHNAKTPREEQYNSTYRVVSVLTVLHQDWLNGKEYCKVKSNKA 126
Dy 294 HEDPEVKFNWYVDGVEVHNAKTPREEQYNSTYRVVSVLTVLHQDWLNGKEYCKVKSNKA 353
Qy 127 LPAPIEKTISKAKGQPREPQVYTLPPSRDELTKNOVSLTCLVKGYPSDIAVWESNGQP 186
Dy 354 LPAPIEKTISKAKGQPREPQVYTLPPSRDELTKNOVSLTCLVKGYPSDIAVWESNGQP 413
Qy 187 ENNYKTTPTPLVLDSDGSFFLYSKLTVDKSRWQOGNVFSCSVMEALHNNHYTKQSLSPCK 246
Dy 414 ENNYKTTPTPLVLDSDGSFFLYSKLTVDKSRWQOGNVFSCSVMEALHNNHYTKQSLSPCK 473
RESULT 11
AAW48650
ID AAW48650 standard; Protein: 652 AA.
```

|||||  
Db 732 PCAPPELLGSPVFLPPPKDFTLMSRTEVTCVVVDVSHEDPEVKFNWYVDGVEVHNA 791  
QY 87 KTRPREEQYNSTYRVVSVLTVLHQDWLNGKEYKCKVSNKALPAPIEKTISKAKGQPREPQ 146  
Db 792 KTRPREEQYNSTYRVVSVLTVLHQDWLNGKEYKCKVSNKALPAPIEKTISKAKGQPREPQ 851  
QY 147 VYTLPSRDELTKNQVSLTCLVKGFYPSDIAVWESNGQPENNYKTTPPVLDSDGSFFLY 206  
Db 852 VYTLPSRDELTKNQVSLTCLVKGFYPSDIAVWESNGQPENNYKTTPPVLDSDGSFFLY 911  
QY 207 SKLTVDKSRWQGNVFCSCVMHEALHNHYTQKSLSLSPGK 246  
Db 912 SKLTVDKSRWQGNVFCSCVMHEALHNHYTQKSLSLSPGK 951  
  
RESULT 8  
AAY92187  
ID AAY92187 standard; protein; 961 AA.  
AC AAY92187;  
XX  
DT 01-AUG-2000 (first entry)  
XX  
DE Human gp130-J-C-gamma-1.  
XX  
KW gp130-J-C-gamma-1; cytokine; antagonist; CNTF; receptor; fusion protein;  
KW cytostatic; immunomodulator; osteopathic.  
XX  
OS Synthetic.  
OS Homo sapiens.  
XX  
FH Key  
FT Protein  
FT 1..619 Location/Qualifiers  
FT /label= gp130  
FT Peptide  
FT 620..621 /note= "Ser-Gly bridge"  
FT Peptide  
FT 622..631 /note= "J-peptide"  
FT Protein  
FT 632..961 /label= IgG1\_constant\_region  
XX  
PN WO200018932-A2.  
XX  
XX  
XX  
PD 06-APR-2000.  
XX  
PF 22-SEP-1999; 99WO-US22045.  
XX  
PR 25-SEP-1998; 98US-0101858.  
PR 19-MAY-1999; 99US-0313942.  
XX  
XX (REG-) REGENERON PHARM INC.  
XX  
XX Stahl N, Yancopoulos GD;  
XX WPI; 2000-293165/25.  
XX  
XX Isolated nucleic acid molecule for treating cytokine-related diseases  
PT or disorders encodes a fusion polypeptide capable of binding a cytokine  
PT to form a nonfunctional complex  
XX  
XX Example 4; Page 7; 152pp; English.  
XX  
XX The invention concerns production of antagonists to any cytokine that  
CC utilizes an alpha specificity determining component, which when combined  
CC with the cytokine, binds to a first beta signal transducing component to  
CC form a non-functional intermediate which then binds to a second beta  
CC signal transducing component causing beta-receptor dimerization, the  
CC soluble alpha specificity determining component of the receptor  
CC (sr-alpha) and the extracellular domain of the first beta signal  
CC transducing component of the cytokine receptor (beta-1) are combined to  
CC form heterodimers (sr-alpha:beta-1) that act as antagonist to the  
CC cytokine by binding the cytokine to form a non-functional complex. The

CC receptor components are shared by cytokines such as the CNTF (ciliary  
CC neurotrophic factor) family of cytokines. The invention provides the  
CC basis for the development of IL-6 antagonists, as they show that if, in  
CC the presence of a ligand, a non-functional intermediate complex,  
CC consisting of the ligand, its alpha receptor and its beta-1 receptor  
CC component, can be formed, it will effectively block the action of the  
CC ligand. Effective antagonists of IL-6 or CNTF consist of heterodimers  
CC of the extracellular domains of the alpha specificity determining  
CC components of their receptors and the extracellular domain of gp130.  
CC The resultant heterodimers, function as high-affinity traps, rendering  
CC the cytokine inaccessible to form a signal transducing complex with the  
CC native membrane-bound forms of their receptor. The nucleic acids and  
CC polypeptides are useful for treating cytokine-related diseases or  
CC disorders such as osteoporosis and primary and secondary effects of  
CC cancer including multiple myeloma or cachexia.  
XX  
SQ Sequence 961 AA;  
Query Match 94.2%; Score 1257; DB 21; Length 961;  
Best Local Similarity 47.9%; Pred. No. 1e-35;  
Matches 236; Conservative 4; Mismatches 6; Indels 247; Gaps 3;  
  
QY 1 DW-----LKA----- 5  
||  
Db 469 DMQOEDGTVHRTYLRGNLAESKCYLITVTPVYADGPGSPESIKAYLKQAPSPGPTVRTK 528  
||  
QY 6 -----FY----- 7  
||  
Db 529 KVGKNEAVLEWDQLPVDVQNGFIRNYTIFRTIIGNETAVNVDSHTEYTLSSLTSDTLY 588  
||  
QY 8 ----- 7  
Db 589 MVRMAAYTDEGGKDGPEFTTTPKFAQGEIESGGQGLTVTVSSASTKGPSVFPFLAPSSKS 648  
QY 8 ----- 7  
Db 649 TSGGTAALGCLVKDYFPEPVTVSWNSGALTSGVHTFPAVLQSSGLYSLSSVTPVSPSSLG 708  
QY 8 -----DKVAEKLKAEFMDKTHCTCPAPPELLGGPSVFLFPPKPKDTLMIS 53  
|| :||: |||||  
Db 709 TQTYICNVNHKPSNTKVDKKVEPKSCDKTHCTCPAPPELLGGPSVFLFPPKPKDTLMIS 768  
QY 54 RTPEVTCVVVDVSHEDPEVKFNWYVDGVEVHNAKTKTPREQYNSTYRVVSVLTVLHQDWL 113  
|||||  
Db 769 RTPEVTCVVVDVSHEDPEVKFNWYVDGVEVHNAKTKTPREQYNSTYRVVSVLTVLHQDWL 828  
QY 114 NGKEYKCKVSNKALPAPIEKTISKAKGQPREPQVYTLPPSRDELTKNQVSLTCLVKGFYP 173  
Db 829 NGKEYKCKVSNKALPAPIEKTISKAKGQPREPQVYTLPPSRDELTKNQVSLTCLVKGFYP 888  
QY 174 SDIAVWESNGQPENNYKTTPPVLDSDGSFFLYSKLTVDKSRWQQGNVFCSCVMHEALHN 233  
Db 889 SDIAVWESNGQPENNYKTTPPVLDSDGSFFLYSKLTVDKSRWQQGNVFCSCVMHEALHN 948  
QY 234 HYTQKSLSLSPGK 246  
|||||  
Db 949 HYTQKSLSLSPGK 961  
  
RESULT 9  
AAB83156  
ID AAB83156 standard; protein; 583 AA.  
XX  
AC AAB83156;  
XX  
DT 02-JUL-2001 (first entry)  
XX  
XX Ganglioside GM2 antibody-related protein #1.  
XX  
KW Ganglioside; GM2; antibody; cytostatic; cytotoxic; cancer.  
XX  
OS Unidentified.  
XX

```
Db 372 RWKSHLQNYTVNATKLTNLTNDRYLATLTVRNLVGKSDAAVLTIPTACDFQATHPVMDLK 431
QY 5 AF-----6
Db 432 APFKNMLWVETTPRESVKYILEWCVLSDKAPCITDWOQEDGTVHRTYLRGNLAESKC 491
QY 7 -----8
Db 492 YLITVTPVYADGPGSPESIKAYLKQAPPSKGTPTVRTKVKGNKAEVLEWDLQPLVDVQNGFI 551
QY 9 -----8
Db 552 RNYTIFYRTIIGNETAVNVDSHSHTYTLSSLTSDTLNVRMAAYTDEGKGDPFTTTP 611
QY 9 -----8
Db 612 KFAQGEIESGASTKGPSVFPPLAPSSKSTSGGTAALGCLVKDYFPFPTVTSWNSGALTSKV 671
QY 9 -----KVAEKLKEAFMDKTHTCP 26
Db 672 HTFPVAVLQSSGLYSLSSVVTVPSSSLGTQTYICNVNHNKPSNTKVDKVEPKSCDKTHTCP 731
QY 27 PCPAPPELLGGPSVFLFPPKPKDTLMISRTPEVTCVVDVSHEDPEVKFNWYVDGVEVHNA 86
Db 732 PCPAPPELLGGPSVFLFPPKPKDTLMISRTPEVTCVVDVSHEDPEVKFNWYVDGVEVHNA 791
QY 87 KTKPREQYNSTYRVVSVLTVLHQDLNKGKEYKCKVSNKALPAPTEKTIISKAKGQPREPQ 146
Db 792 KTKPREQYNSTYRVVSVLTVLHQDLNKGKEYKCKVSNKALPAPTEKTIISKAKGQPREPQ 851
QY 147 VYTLPPSRDELTKNQVSLTCLVKGPYPSDIAVWESNQPENNYKTPPVLDSDGSFFLY 206
Db 852 VYTLPPSRDELTKNQVSLTCLVKGPYPSDIAVWESNQPENNYKTPPVLDSDGSFFLY 911
QY 207 SKLITVDKSRWQQGNVFSCSVMHEALHNHYTQKSLSLSPGK 246
Db 912 SKLITVDKSRWQQGNVFSCSVMHEALHNHYTQKSLSLSPGK 951
```

```
RESULT 7
AA92186
ID AA92186 standard; protein: 951 AA.
XX
AC AA92186;
XX
DT 01-AUG-2000 (first entry)
XX
DE Human gp130-C-gamma-1.
XX
KW gp130-C-gamma-1; cytokine; antagonist; CNTF; receptor; fusion protein;
KW cytosolic; immunomodulator; osteopathic.
XX
OS Synthetic.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT Protein 1..619
FT /label= gp130
FT Peptide 620..621
FT /note= "Ser-Gly bridge"
FT Protein 622..951
FT /label= IgG1_constant_region
XX
XX WO200018932-A2.
XX
XX 06-APR-2000..
XX
XX 22-SEP-1999; 99WO-US22045.
XX
XX 25-SEP-1998; 98US-0101858.
XX
XX 19-MAY-1999; 99US-0313942.
XX
```

```
PA (REGE-) REGENERON PHARM INC.
XX Stahl N, Yancopoulos GD;
DR WPI: 2000-293165/25.
XX Isolated nucleic acid molecule for treating cytokine-related diseases
PT or disorders encodes a fusion polypeptide capable of binding a cytokine
PT to form a nonfunctional complex
XX Example 4; Fig 9; 152pp; English.
PS
XX The invention concerns production of antagonists to any cytokine that
CC utilizes an alpha specificity determining component, which when combined
CC with the cytokine, binds to a first beta signal transducing component to
CC form a non-functional intermediate which then binds to a second beta
CC signal transducing component causing beta-receptor dimerization, the
CC soluble alpha specificity determining component of the receptor
CC (SR-alpha) and the extracellular domain of the first beta signal
CC transducing component of the cytokine receptor (beta-1) are combined to
CC form heterodimers (SR-alpha:beta-1) that act as antagonist to the
CC cytokine by binding the cytokine to form a non-functional complex. The
CC receptor components are shared by cytokines such as the CNTF (ciliary
CC neurotrophic factor) family of cytokines. The invention provides the
CC basis for the development of IL-6 antagonists, as they show that if, in
CC the presence of a ligand, a non-functional intermediate complex,
CC consisting of the ligand, its alpha receptor and its beta-1 receptor
CC component, can be formed, it will effectively block the action of the
CC ligand. Effective antagonists of IL-6 or CNTF consist of heterodimers
CC of the extracellular domains of the alpha specificity determining
CC components of their receptors and the extracellular domain of gp130.
CC The resultant heterodimers, function as high-affinity traps, rendering
CC the cytokine inaccessible to form a signal transducing complex with the
CC native membrane-bound forms of their receptor. The nucleic acids and
CC polypeptides are useful for treating cytokine-related diseases or
CC disorders such as osteoporosis and primary and secondary effects of
CC cancer including multiple myeloma or cachexia.
XX
SQ Sequence 951 AA;
```

```
Query Match 94.2%; Score 1257; DB 21; Length 951;
Best Local Similarity 37.0%; Pred. No. 1e-35; 5; Indels 394; Gaps 3;
Matches 237; Conservative 4; Mismatches 5;
QY 1 DW-----2
Db 312 DWSEASGITYEDRPSKAPSFWKIDPSHTQGYRTVQLVWKTLPPEANGKILDYEVTLT 371
QY 3 -----LK 4
Db 372 RWKSHLQNYTVNATKLTNLTNDRYLATLTVRNLVGKSDAAVLTIPTACDFQATHPVMDLK 431
QY 5 AF-----6
Db 432 APFKNMLWVETTPRESVKYILEWCVLSDKAPCITDWOQEDGTVHRTYLRGNLAESKC 491
QY 7 -----YD-----8
Db 492 YLITVTPVYADGPGSPESIKAYLKQAPPSKGTPTVRTKVKGNKAEVLEWDLQPLVDVQNGFI 551
QY 9 -----8
Db 552 RNYTIFYRTIIGNETAVNVDSHSHTYTLSSLTSDTLNVRMAAYTDEGKGDPFTTTP 611
QY 9 -----8
Db 612 KFAQGEIESGASTKGPSVFPPLAPSSKSTSGGTAALGCLVKDYFPFPTVTSWNSGALTSKV 671
QY 9 -----KVAEKLKEAFMDKTHTCP 26
Db 672 HTFPVAVLQSSGLYSLSSVVTVPSSSLGTQTYICNVNHNKPSNTKVDKVEPKSCDKTHTCP 731
QY 27 PCPAPPELLGGPSVFLFPPKPKDTLMISRTPEVTCVVDVSHEDPEVKFNWYVDGVEVHNA 86
```

CC been constructed, which may be useful for treating neurological  
CC disorders. The ephrin fusion proteins are preferably capable of binding  
CC to Elk receptor and are especially Efl-6 antagonists. The fusion proteins  
CC were constructed after it was demonstrated that similar improved  
CC activities could be achieved using Tie-2 receptor ligands.  
CC Angiopoietin-1 (Ang-1) is a Tie-2 receptor ligand and is an agonist for  
CC Tie-2, whereas angiopoietin-2 (Ang-2) is a naturally occurring antagonist  
CC of the Tie-2 receptor. The fibrinogen domains (FD) of Ang-1 and Ang-2 are  
CC the receptor-binding domains and dimerized versions, of e.g. Ang-1-FD-FC  
CC (Ang-1 fibrinogen domain fused to an FC domain), can bind to the Tie-2  
CC receptor with much higher affinity than monomeric Ang-1-FD (dimerization  
CC occurs between the FC components of adjacent molecules). However,  
CC Ang-1-FD-FC is not able to induce phosphorylation (activate) the Tie-2  
CC receptor on endothelial cells unless it is further clustered with goat  
CC anti-human FC antibodies. The novel fusion proteins, mutant versions of  
CC Ang-1-FD and Ang-2-FD, were designed that were intrinsically more highly  
CC clustered. Tie-2 agonist fusion proteins may be used as haematopoietic  
CC factors. Tie-2 receptor antagonist fusion proteins may be used to  
CC diagnose or treat, e.g. myeloproliferative or other proliferative  
CC disorders of blood forming organs, e.g. thrombocythemias, polycythemia  
CC and leukemias.

XX Sequence 683 AA;

Query Match 94.2%; Score 1257; DB 21; Length 683;  
Best Local Similarity 35.4%; Pred. No. 4 6e-36;  
Matches 238; Conservative 3; Mismatches 4; Indels 427; Gaps 4;

QY 2 WLKA-----FVDKVAEL-----14  
|||  
|||

Db 12 WLVMVWALRLATPLAKNLEPVSWSLNPFLSGKGLVIVPKIGDKLDIICPRAEAGR 71  
|||  
|||

QY 15 -----14

Db 72 PYEYKLYLVPRQAAACSTVLDPNVLVTCNRPQEIRFTIKQEPSPNMYGLFEXKHHD 131

QY 15 -----17  
|||  
|||

Db 132 YIITSTNGSLEGNREGGVCRTMKIIMKVGQDPNAVTPQLTTSRPSREADNTVMK 191  
|||  
|||

QY 18 -----17

Db 192 ATQAPGSGSLGSDGKHETVNOEKGSGPGASGGSDPGDFNKGPKGNLEPVSWSL 251

QY 18 -----17

Db 252 NPKFLSGGLVYIPKIGDKLDIICPRAEAGRPEYKLYLVPRQAAACSTVLDPNVLT 311

QY 18 -----FM-----19  
:|

Db 312 CNRPEQIRFTIKQEPSPNMYGLEFKKHDDYIITSTNGSLEGNREGGVCRTMKI 371

QY 20 -----19

Db 372 IMKVGQDPNAVTPQLTTSRPSKEADNTVMATQAPGSRGSLGSDGKHETVNOEKGSGP 431

QY 20 -----DKTHTCPCAPPELLGGPSVFLFPKPKDTLMISR 54  
|||||

Db 432 GASGGSGDPPGDFNKGPKGPKSCDKTHTCPCAPPELLGGPSVFLFPKPKDTLMISR 491  
|||||

QY 55 TPEVTCVVVDVSHEDPEVKFNMYVDGVEVHNATKPREQYNSTYRVVSVLTVLHQDWLN 114  
|||||

Db 492 TPEVTCVVVDVSHEDPEVKFNMYVDGVEVHNATKPREQYNSTYRVVSVLTVLHQDWLN 551  
|||||

QY 115 GREYKCKVSNKALPAPIETKISKAKQPREPVYITLPPSRDELTKNQVSLTCLVKGFYPS 174  
|||||

Db 552 GREYKCKVSNKALPAPIETKISKAKQPREPVYITLPPSRDELTKNQVSLTCLVKGFYPS 611  
|||||

QY 175 DLAVERESNGQPNVYKTTTPVLDSDGSEFLYSLKLVTDKSRWQOGNVFSCSVNHEALHNH 234  
|||||

Db 612 DLAVERESNGQPNVYKTTTPVLDSDGSEFLYSLKLVTDKSRWQOGNVFSCSVNHEALHNH 671  
|||||

QY 235 YTKQKSLSLSPGK 246

Db 672 YTKQKSLSLSPGK 683  
|||||

# RESULT 6

AAW70798

XX AAW70798 standard; protein; 951 AA.

AC AAW70798;

XX 03-FEB-1999 (first entry)

XX Human gpl30-C-gamma-1 amino acid sequence.

XX gpl30; cytokine antagonist; interleukin; gamma-interferon;

XX granulocyte macrophage colony-stimulating factor; J peptide;

XX transforming growth factor-beta.

OS Synthetic.

OS Homo sapiens.

XX Key Location/Qualifiers

FT Protein 1..619

FT Protein /note= "human gpl30"

FT Protein 662..651

XX /note= "from the constant region of human IgG1"

PN US5844099-A.

XX 01-DEC-1998.

XX 27-NOV-1995; 95US-0563105.

XX 27-NOV-1995; 95US-0563105.

XX 20-OCT-1993; 93US-0140222.

XX (REGE-) REGENERON PHARM INC.

XX Economides A, Stahl N, Yancopoulos GD;

XX WPI; 1999-044669/04.

XX Cytokine antagonists - comprising extracellular domains of

XX specificity-determining and signal-transducing components of

XX cytokine receptor

PS Example 4; Fig 9A-B; 46pp; English.

XX The present sequence represents the amino acid sequence of human

XX gpl30-C-gamma-1. The protein is used in the course of the invention.

XX The specification describes cytokine antagonists comprising only the

XX extracellular domain of the specificity-determining component of a

XX the cytokine receptor and the extracellular domain of a

XX signal-transducing component of the cytokine receptor. The cytokine

XX is an interleukin (IL-1, IL-2, IL-3, IL-4, IL-5 or IL-15),

XX granulocyte macrophage colony-stimulating factor (GM-CSF),

XX gamma-interferon or transforming growth factor-beta (TGF-beta). The

XX antagonist is capable of binding the cytokine to form a nonfunctional

XX complex. The compounds have therapeutic activity as cytokine antagonists

XX and can also be used in assays for identifying novel agonists and

XX antagonists of cytokines.

XX Sequence 951 AA;

Query Match 94.2%; Score 1257; DB 20; Length 951;

Best Local Similarity 37.0%; Pred. No. 1e-35;

Matches 237; Conservative 4; Mismatches 5; Indels 394; Gaps 3;

QY 1 DW-----2

Db 312 DWSEASGTYEDRPSKAPSFYWKIDPSHTQGYRTVOLVWKTLPPFEANGKILDYEVTLT 371

QY 3 -----LK 4



XX DE Protein sequence of primatised form of the heavy chain of 16C10 antibody.

XX KW Human; macaque monkey; light chain; primatised antibody; 16C10 antibody;

KW neuroprotective; apoptosis inducer; allergy; CD28 receptor antagonist;

KW B7.1 antigen; CD80; B7.2 antigen; CD86; B cell cancer; metastasis;

KW tumour; B cell lymphoma; B cell leukaemia; autoimmune disease;

KW graft-vs-host disease; immunosuppression; organ rejection;

KW interleukin-2; IL-2; mutant; mutein.

XX KW Chimeric - Homo sapiens.

OS Chimeric - Macaca Sp.

OS Synthetic.

XX PN WO200189567-A1.

XX PD 29-NOV-2001.

XX PF 22-MAY-2001; 2001WO-US16364.

XX PR 22-MAY-2000; 2000US-0576424.

XX PA (IDEC-) IDEC PHARM CORP.

XX PI Anderson DR, Hanna N, Brams P;

XX DR WPI; 2002-089895/12.

XX DR N-PSDB; AAS17247.

XX PT Use of monoclonal antibody which specifically binds to B7.1 antigen

PT CD80 and/or B7.2 antigen CD86 for inducing apoptosis of B7+ cells,

PT treating cancer, graft-vs-host disease and autoimmune disease as

PT allergy

XX PS Example 8; Fig 5b; 89pp; English.

XX CC The present invention relates to a new use of a monoclonal antibody

CC which specifically binds to B7.1 antigen (CD80) and/or B7.2 antigen

CC (CD86) for inducing the apoptosis of B7+ cells. The invention is

CC useful for treating diseases such as B cell cancer, lymphoma, a

CC cancer where B cells promote the growth and/or metastasis of tumours,

CC B cell lymphoma, B cell leukaemia, and autoimmune diseases such as

CC idiopathic thrombocytopenia purpura, systemic lupus, erythematosis,

CC type 1 diabetes mellitus, rheumatoid arthritis, psoriasis, aplastic

CC anaemia, inflammatory bile disease, allergy, multiple sclerosis

CC or graft-vs-host disease. The antibody is useful for immunosuppression

CC in a human or animal and for treating or preventing resistance to or

CC rejection of transplanted organ or tissue for treating proliferative

CC and hyperproliferative diseases, for treating reversible obstructive

CC airways disease, intestinal inflammations and allergies e.g. Crohn's

CC disease and ulcerative colitis, food-related allergies e.g. migraine,

CC rhinitis and eczema, and other types of allergies. The present protein

CC sequence represents the heavy chain of 16C10, a primatised antibody

CC used in the invention to induce apoptosis and inhibit production of

XX interleukin-2 (IL-2).

XX SQ Sequence 476 AA;

Query Match 94.3%; Score 1258; DB 23; Length 476;

Best Local Similarity 56.7%; Pred. No. 1.8e-36;

Matches 233; Conservative 5; Mismatches 8; Indels 165; Gaps 1;

QY 1 DMLKAFYD----- 8

DB 66 EWIGSFYSSGNTYNNPSLKSQVTTSTDTKQNFSLKNSMTAATVYICVRDLFSV 125

QY 9 ----- 8

DB 126 GMVYNNWFVWGPGLVTVSSASTKGPVFLAPSSKSTSGTAALGLVDFPEPVTV 185

QY 9 -----KVAELK 15

DB 186 SWNSGALTSGVHTFFPAVLQSLGSLSSVTVVPSSSLGTQTYICNVNHPKSNKTKVDRKAE 245

QY 16 EAFMDKTHTCPCPAPELLGGPSVFLFPPKPKDILMISRTPEVTCVVVDVSHEDPEVKFN 75

DB 246 PKSCDKTHTCPCPAPELLGGPSVFLFPPKPKDILMISRTPEVTCVVVDVSHEDPEVKFN 305

QY 76 WYDGVVEVHNAKTPREQYNSTYRVVSVLTVLHQDWLNGKEYCKVSNKALPAPIEKTI 135

DB 306 WYDGVVEVHNAKTPREQYNSTYRVVSVLTVLHQDWLNGKEYCKVSNKALPAPIEKTI 365

QY 136 SKAGQPREPOVYITLPPSRDELTKNOVSLTCLVKGFYPSDIAVEWESNGQPENNYKTTTP 195

DB 366 SKAGQPREPOVYITLPPSRDELTKNOVSLTCLVKGFYPSDIAVEWESNGQPENNYKTTTP 425

QY 196 VLSDGSFFLYSKLTVDKSRWQQGNVFCSCVMHEALHNYHQKSLSPGK 246

DB 426 VLSDGSFFLYSKLTVDKSRWQQGNVFCSCVMHEALHNYHQKSLSPGK 476

RESULT 5

AAY96781

ID AAY96781 standard; Protein; 683 AA.

XX AC AAY96781;

XX DT 26-SEP-2000 (first entry)

XX XX Ephrin-B1-Ephrin-B1-FC fusion protein.

DE Ephrin-B1; ELK receptor; ligand; dimer; Fc domain; fusion protein;

KW Efl-6 antagonist; neurological.

XX OS Chimeric - Homo sapiens.

OS Chimeric - Synthetic.

XX Key Location/Qualifiers

FT Peptide 1..29

FT /label= Signal\_peptide

FT /note= "derived from ELK-L Ectodomain 1"

FT Protein 30..237

FT /label= ELK-L\_ectodomain\_1

FT Peptide 238..240

FT /label= Bridging\_peptide

FT Protein 241..448

FT /label= ELK-L\_Ectodomain\_2

FT Peptide 449..451

FT /label= Bridging\_peptide

FT Region 452..683

FT /label= Human\_IgG1\_Fc\_region

XX WO200037642-A1.

XX 29-JUN-2000.

XX 23-DEC-1999; 99WO-US30900.

XX 23-DEC-1998; 98US-0113387.

XX (REG-) REGENERON PHARM INC.

XX Davis SJ, Gale NW, Yancopoulos GD, Stahl N;

XX WPI; 2000-442670/38.

DR N-PSDB; AAA51345.

XX Polynucleotide encoding a fusion polypeptide, useful for promoting

PT differential function and influencing phenotype, comprises two subunits

PT containing at least one copy of the receptor binding domain of a ligand

XX Example 12; Fig 14A-E; 97pp; English.

PS Production of homogenous forms of clustered ligands is broadly applicable

CC to improve the affinity and/or increase the activity of a ligand as

CC compared to the native form of the ligand. Ephrin fusion proteins have

CC B7.1 antigen monoclonal antibody 16C10. Cloned 16C10 light and heavy variable genes (see also AAT62512 and AAT62513) are inserted into an expression vector (pref. NEOSPLA) which contains human light and heavy chain constant region genes to allow prodn. of the primatised antibody in e.g. CHO cells. Primatised 7C10 and 7B6 anti-B7.1 antibodies have also been produced. (see also AAW01817-20). The primatised antibodies inhibit the B7:CD28 pathway, making them useful immunosuppressants for the treatment of autoimmune disorders and graft-versus-host disease.

XX SQ Sequence 476 AA;

Query Match 94.3%; Score 1258; DB 18; Length 476;  
Best Local Similarity 56.7%; Pred. No. 1.8e-36;  
Matches 233; Conservative 5; Mismatches 8; Indels 165; Gaps 1;

QY 1 DWLKAFYD----- 8  
: : : |||

Db 66 EWIGSFYSSSGNTYYNPSLKSQVTTSTDTSKNQFSLKLNMTAADTAVYYCVRDLFSVY 125  
: : : |||

QY 9 ----- 8

Db 126 GMVYNNWFDVWGPGVLVTVSSASTKGPSVFPLAPSSKSTSGGTAALGCLVKDYFPEPTV 185

QY 9 -----KVAEKLK 15  
: : : |||

Db 186 SNNSGALTSGVHTFPAVLQSSGLYSLSSVTVTPSSSLGTQTYICNVNHHKPSNTKVDKKA 245  
: : : |||

QY 16 EAFMDKTHTCPCPAPELLGGPSVFLFPPKPKDTLMISRTPEVTCVVDVSHEDPEVKFN 75  
|||||

Db 246 PKSCDKTHTCPCPAPELLGGPSVFLFPPKPKDTLMISRTPEVTCVVDVSHEDPEVKFN 305  
|||||

QY 76 WYVDGVEVHNATKPREEQYNSTYRVSVSLTVLHODWLNKGEYCKKVSNNKALPAPIEKTI 135  
|||||

Db 306 WYVDGVEVHNATKPREEQYNSTYRVSVSLTVLHODWLNKGEYCKKVSNNKALPAPIEKTI 365  
|||||

QY 136 SKAKGQPREPQVYTLPPSRDELTKNQVSLTCLVKGFYPSDIAVEVESNGQPENNYKTTP 195  
|||||

Db 366 SKAKGQPREPQVYTLPPSRDELTKNQVSLTCLVKGFYPSDIAVEVESNGQPENNYKTTP 425  
|||||

QY 196 VLDSGFSFLYSLKLTVDKSRWQGNVFCVSMVHEALHNHYTQKSLSLSPGK 246  
|||||

Db 426 VLDSGFSFLYSLKLTVDKSRWQGNVFCVSMVHEALHNHYTQKSLSLSPGK 476  
|||||

RESULT 3  
AAW63765  
ID AAW63765 standard; Protein; 476 AA.  
XX AC AAW63765;  
XX DT 29-SEP-1998 (first entry)  
XX DE Macaque primatized 16C10 heavy chain protein.  
XX KW Monoclonal antibody; Mab; macaque; heavy chain; primate; antigen; CD80;  
KW CD86; inhibitor; immunosuppressant; treatment; autoimmune disease; IL-2;  
KW T cell/B cell interaction; tumour; inflammation; imaging agent; vaccine;  
KW immunogen; anti-idiotype reagent; interleukin-2; IgG; immunoglobulin G;  
KW T cell proliferation.  
XX OS Macaca fascicularis.  
XX PN W09B19706-A1.  
XX PD 14-MAY-1998.  
XX PF 29-OCT-1997; 97WO-US19906.  
XX PR 08-NOV-1996; 96US-0746361.  
XX PA (IDEC-) IDEC PHARM CORP.  
XX DT

PI Anderson DR, Brams P, Hanna N;  
XX WPI: 1998-286601/25.  
DR N-PSDB; AAV35489.  
XX New monoclonal antibodies specific for B7.1 or B7.2 antigens and inhibiting binding to CD28 - useful as specific immunosuppressants for treating diseases that involve interactions between T and B cells, e.g. graft rejection or tumours

PT PT  
PT cells, e.g. graft rejection or tumours

XX PS Example 7; Fig 5b; 87pp; English.

XX This sequence represents a primatized form of the antibody 16C10 heavy chain from macaque. This sequence is used in a method which studies new monoclonal antibodies (MAB's) that bind selectively to B7.1 (CD80) or to B7.2 (CD86) antigens and inhibits binding of these antigens to CD28. Such MAB's are specific immunosuppressants for treatment of diseases involving T cell/B cell interactions, particularly autoimmune disease, specifically idiopathic thrombocytopenia purpura, systemic lupus erythematosus, type I diabetes mellitus, rheumatoid arthritis, psoriasis, aplastic anaemia, inflammatory bowel disease, allergy and multiple sclerosis, graft vs. host diseases, B cell lymphoma, infections (including by human immune deficiency virus) or inflammatory disease and tumours. Optionally the MAB can be conjugated to a drug or toxin. MAB's, or their fragments, can also be used as imaging agents and as vaccines or immunogens to develop anti-idiotype reagents. MAB's are optionally combined with other proteins or small molecule immunosuppressants. Blocking B7/CD28 interactions induces long-term, antigen-specific immunosuppression, i.e. it inhibits production of interleukin-2 (IL-2), T cell proliferation and antigen-specific immunoglobulin G (IgG) responses.

XX SQ Sequence 476 AA;

Query Match 94.3%; Score 1258; DB 19; Length 476;  
Best Local Similarity 56.7%; Pred. No. 1.8e-36;  
Matches 233; Conservative 5; Mismatches 8; Indels 165; Gaps 1;

QY 1 DWLKAFYD----- 8  
: : : |||

Db 66 EWIGSFYSSSGNTYYNPSLKSQVTTSTDTSKNQFSLKLNMTAADTAVYYCVRDLFSVY 125  
: : : |||

QY 9 ----- 8

Db 126 GMVYNNWFDVWGPGVLVTVSSASTKGPSVFPLAPSSKSTSGGTAALGCLVKDYFPEPTV 185

QY 9 -----KVAEKLK 15  
: : : |||

Db 186 SNNSGALTSGVHTFPAVLQSSGLYSLSSVTVTPSSSLGTQTYICNVNHHKPSNTKVDKKA 245  
: : : |||

QY 16 EAFMDKTHTCPCPAPELLGGPSVFLFPPKPKDTLMISRTPEVTCVVDVSHEDPEVKFN 75  
|||||

Db 246 PKSCDKTHTCPCPAPELLGGPSVFLFPPKPKDTLMISRTPEVTCVVDVSHEDPEVKFN 305  
|||||

QY 76 WYVDGVEVHNATKPREEQYNSTYRVSVSLTVLHODWLNKGEYCKKVSNNKALPAPIEKTI 135  
|||||

Db 306 WYVDGVEVHNATKPREEQYNSTYRVSVSLTVLHODWLNKGEYCKKVSNNKALPAPIEKTI 365  
|||||

QY 136 SKAKGQPREPQVYTLPPSRDELTKNQVSLTCLVKGFYPSDIAVEVESNGQPENNYKTTP 195  
|||||

Db 366 SKAKGQPREPQVYTLPPSRDELTKNQVSLTCLVKGFYPSDIAVEVESNGQPENNYKTTP 425  
|||||

QY 196 VLDSGFSFLYSLKLTVDKSRWQGNVFCVSMVHEALHNHYTQKSLSLSPGK 246  
|||||

Db 426 VLDSGFSFLYSLKLTVDKSRWQGNVFCVSMVHEALHNHYTQKSLSLSPGK 476  
|||||

RESULT 4  
AAU11646  
ID AAU11646 standard; Protein; 476 AA.  
XX AC AAU11646;  
XX DT 12-MAR-2002 (first entry)

PI	Kainoh M, Tanaka T;
XX	
DR	WPI; 1998-427881/36.
DR	N-PSDB; AAV33772.
XX	
PT	Integrin-immunoglobulin chimeric protein heterodimer complexes as
PT	platelet substitutes contain the alpha and beta integrin chains
PT	associated in stable state and bind to extracellular matrix in the
PT	presence of plasma components
XX	
PS	Claim 7; Pages 42-50; 87pp; Japanese.
XX	
CC	This represents an integrin alpha-4 chain. The invention provides
CC	integrin-immunoglobulin chimeric protein heterodimer complexes that
CC	comprise an integrin alpha or beta chain associated with an
CC	immunoglobulin light or heavy chain. These chimeric proteins form
CC	heterodimer complexes, in particular with a chimeric protein containing
CC	an integrin alpha chain and an immunoglobulin chain with a chimeric
CC	protein containing an integrin beta chain and an immunoglobulin chain;
CC	the immunoglobulin chain in each case may be a heavy chain, or one of the
CC	two may be a light chain. The integrin alpha chain is preferably alpha 4
CC	or alpha 2 and the integrin beta chain is preferably beta 1. Animal cells
CC	transformed with vectors containing the DNA coding for the above chimeric
CC	proteins can be used in the preparation of the chimeric proteins and
CC	their heterodimer complexes. The heterodimer complexes, which are useful
CC	for testing potential promoters and inhibitors of the binding of
CC	integrins to their ligands, function as blood platelet substitutes and
CC	hemostatics and as diagnostic agents.
XX	
SQ	Sequence 1218 AA;
	Query Match 94.5%; Score 1261; DB 19; Length 1218;
	Best Local Similarity 24.2%; Pred. No. 1.3e-35;
	Matches 236; Conservative 4; Mismatches 3; Indels 734; Gaps 3;
QY	4 KAFYDK-----VAE 12
Db	242 KAFLDKNQVKGSLGYSVGAGHFRSHTTEVVGAPQHGIKAYIFSIDEKEINLH 301
QY	13 KLK-----15
	:::
Db	302 EMKGKKLGSYFGASCAVDLNADGFSDLVGAPMQSTIREGRVFVYINSGGAYNAME 361
QY	16 -----15
Db	362 TLNVGSKYARFCESIVNLGDINDGFEDVAIGAQPDDLOGAIYYINGRADGISSTFS 421
QY	16 -----15
Db	422 QRIEGLQISLSMFGQSISQIDADNNGYVDVAVGAFRSDSAVLRLRPVVIVDASLSH 481
QY	16 -----15
Db	482 PESYNRTKFDVCENGWP SVCIDLTLCF SYKGKEVPGYIVLFYNKSLDNRKRAESPREFY 541
QY	16 -----EAFM-----19
	:
Db	542 SSNSTSDVTIGSIQVSSREANCRTHQAFMRKDVRDLTPIQIEAAYHLGHVVISKRSTEE 601
QY	20 -----19
Db	602 FPLQLIQQKKEKDIMKKTINFARFCAHENSADLQVSAKIGIFLKPHENKTYLAVGSMK 661
QY	20 -----19
Db	662 TLMNLVSLFNAGDAYETTLHKLVGLVLIKFILKELEEKQINCEVTNSGVVQLDCSIGY 721
QY	20 -----19
Db	722 IYVDHLSRDISFLLDVSSLRAEEDLSITVHATCENEEMDNKLHSRVTVAIPLUYEVK 781
QY	20 -----19

D	b	782	LTVHGFVNPTSFVYSGNDENEPETCMVEKMNLTFFHVINTGNSMAPNVSVIEIWPNSFSFQ	841
Q	y	20	-----	19
D	b	842	TDKLFNILDYQTTTGCEHFENYQRVCALBQKSAMOTLAKIVRFLSKTDKRLLYCICKADP	901
Q	y	20	-----	19
D	b	902	HCLFNLCNFGKMSGKEASVHIQLEGRPSILEMDETSALKFEIRATGPFPNPRIEVLNK	961
Q	y	20	-----DKTHTCPPCPAPPELLGGPSVFLFPPPKPDT	49
D	b	962	DENVAHVLLGLELHQRPKYFTDPEPKSCDKTHTCCPCPAPELLGGPSVFLFPPPKPDT	1021
Q	y	50	LMISRTPEVTCCVVVDVSHEDPEVKFNWYVDGEVHNNAKTTPREEQYNSTYRVVSVLTVLH	109
D	b	1022	LMISRTPEVTCCVVVDVSHEDPEVKFNWYVDGEVHNNAKTTPREEQYNSTYRVVSVLTVLH	1081
Q	y	110	QDWLNGKEYKCYSNKAIPAPIEKTTISKAGQPREPOVYTLPPSRDELTKNQVSLTCLVK	161
D	b	1082	QDWLNGKEYKCYSNKAIPAPIEKTTISKAGQPREPOVYTLPPSRDELTKNQVSLTCLVK	1141
Q	y	170	GYPDSIAVSWESNGOPENNKKTPPVLDSDGSFFLYSKLTVDKSRWQQGNVFSCSVMHE	229
D	b	1142	GYPDSIAVSWESNGOPENNKKTPPVLDSDGSFFLYSKLTVDKSRWQQGNVFSCSVMHE	1201
Q	y	230	ALNNHYTKQSLSLSPGK	246
D	b	1202	ALNNHYTKQSLSLSPGK	1218
<b>RESULT 2</b>				
AAW01822				
ID	AAW01822 standard; Protein; 476 AA.			
AC	AAW01822;			
DT	25-MAY-1997 (first entry)			
XX	Primitised anti-human B7.1 antigen antibody 16C10 heavy chain.			
KW	Monoclonal antibody; cynomolgus monkey; macaque; 16C10;			
KW	primatised antibody; B7 antigen; CD28; immunosuppressive;			
KW	autoimmune disease; idiopathic thrombocytopoenia purpura;			
KW	systemic lupus erythematosus; rheumatoid arthritis; psoriasis;			
KW	type 1 diabetes mellitus; graft versus host disease;			
XX	hetero-hybridoma; transfectoma.			
OS	Chimeric Macaca cynomolgus;			
OS	Chimeric Homo sapiens.			
PN	W09640878-A1.			
PD	19-DEC-1996.			
PF	06-JUN-1996; 96WO-US10053.			
PR	07-JUN-1995; 95US-0487550.			
PA	(IDEC-) IDEC PHARM CORP.			
PI	Anderson DR, Brans P, Hanna N, Shestowsky WS;			
XX	WPI; 1997-108638/10.			
XX	N-PSDB; AATG2513.			
PT	Monkey monoclonal antibody binding human B7.1 or B7.2 antigen -			
XX	useful for treating autoimmune disease or graft-versus-host disease			
PS	Claim 14; Fig 10B; 81pp; English.			
CC	2 Polypeptides (AAW01821 and AAW01822) respectively comprise primatises			
CC	forms of the light and heavy chains of cynomolgus monkey anti-human			

GenCore version 5.1.4\_p5\_4578  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: April 21, 2003, 10:42:21 ; Search time 28.9412 Seconds  
(without alignments)  
1132.630 Million cell updates/sec

Title: 7LINK2

Perfect score: 1334

Sequence: 1 DWLKAFYDKVAEKLKAFMD.....MHEALHNYTKLSLSLSPGK 246

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.0

Searched: 908470 seqs, 133250620 residues

Total number of hits satisfying chosen parameters: 908470

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : A\_Geneseq\_101002.\*

- 1: /SIDS2/gcgdata/geneseq/geneseqp-emb1/AA1980.DAT.\*
- 2: /SIDS2/gcgdata/geneseq/geneseqp-emb1/AA1981.DAT.\*
- 3: /SIDS2/gcgdata/geneseq/geneseqp-emb1/AA1982.DAT.\*
- 4: /SIDS2/gcgdata/geneseq/geneseqp-emb1/AA1983.DAT.\*
- 5: /SIDS2/gcgdata/geneseq/geneseqp-emb1/AA1984.DAT.\*
- 6: /SIDS2/gcgdata/geneseq/geneseqp-emb1/AA1985.DAT.\*
- 7: /SIDS2/gcgdata/geneseq/geneseqp-emb1/AA1986.DAT.\*
- 8: /SIDS2/gcgdata/geneseq/geneseqp-emb1/AA1987.DAT.\*
- 9: /SIDS2/gcgdata/geneseq/geneseqp-emb1/AA1988.DAT.\*
- 10: /SIDS2/gcgdata/geneseq/geneseqp-emb1/AA1989.DAT.\*
- 11: /SIDS2/gcgdata/geneseq/geneseqp-emb1/AA1990.DAT.\*
- 12: /SIDS2/gcgdata/geneseq/geneseqp-emb1/AA1991.DAT.\*
- 13: /SIDS2/gcgdata/geneseq/geneseqp-emb1/AA1992.DAT.\*
- 14: /SIDS2/gcgdata/geneseq/geneseqp-emb1/AA1993.DAT.\*
- 15: /SIDS2/gcgdata/geneseq/geneseqp-emb1/AA1994.DAT.\*
- 16: /SIDS2/gcgdata/geneseq/geneseqp-emb1/AA1995.DAT.\*
- 17: /SIDS2/gcgdata/geneseq/geneseqp-emb1/AA1996.DAT.\*
- 18: /SIDS2/gcgdata/geneseq/geneseqp-emb1/AA1997.DAT.\*
- 19: /SIDS2/gcgdata/geneseq/geneseqp-emb1/AA1998.DAT.\*
- 20: /SIDS2/gcgdata/geneseq/geneseqp-emb1/AA1999.DAT.\*
- 21: /SIDS2/gcgdata/geneseq/geneseqp-emb1/AA2000.DAT.\*
- 22: /SIDS2/gcgdata/geneseq/geneseqp-emb1/AA2001.DAT.\*
- 23: /SIDS2/gcgdata/geneseq/geneseqp-emb1/AA2002.DAT.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1261	94.5	1218	19	AAW70539
2	1258	94.3	476	18	AAW01822
3	1258	94.3	476	19	AAW63765
4	1258	94.3	476	23	AAU11646
5	1257	94.2	683	21	AAAY96781
6	1257	94.2	951	20	AAW70798
7	1257	94.2	951	21	AAAY92186
8	1257	94.2	961	21	AAAY92187
9	1255	94.1	583	22	AAAB83156
10	1254	94.0	473	22	AAAG64473
					Integrin alpha-4 c
					Primate anti-hu
					Macaque primatized
					Protein sequence o
					Ephrin-B1-Ephrin-B
					Human gp130-C-gamm
					Human gp130-J-C-ga
					Ganglioside GM2 an
					Human type antihum

11	1254	94.0	652	19	AAW48650	Heavy chain of hma
12	1254	94.0	690	21	AAAY92195	Human IL-6R-alpha-
13	1254	94.0	859	20	AAW70796	Human gp130-Fc-His
14	1254	94.0	859	21	AAAY92184	Human gp130-Fc-His
15	1254	94.0	1158	21	AAAY92205	Fusion polypeptide
16	1254	94.0	1168	21	AAAY92204	Fusion polypeptide
17	1253	93.9	471	21	AAAY92204	HUMAN OCR10-Fc fus
18	1253	93.9	475	13	AAAR20057	Heavy chain of 3D6
19	1253	93.9	475	18	AAAR20057	Human anti-RSV mon
20	1253	93.9	475	18	AAAG1639	Human anti-RSV mon
21	1253	93.9	475	22	AAAG3640	Amino acid sequenc
22	1253	93.9	567	23	AAE13733	Human zaiphalr/1g
23	1253	93.9	754	21	AAAB11691	Human secreted K10
24	1253	93.9	787	21	AAAB11693	Human secreted K10
25	1253	93.9	1367	19	AAW70542	Integrin alpha-2 c
26	1252	93.9	449	14	AAAR43339	Completely humanis
27	1252	93.9	449	19	AAW49816	Amino acid sequenc
28	1252	93.9	475	17	AAAR93553	Monoclonal antibod
29	1251	93.8	595	20	AAW86003	Anti-574 single ch
30	1251	93.8	462	21	AAAB29408	Human monoclonal a
31	1251	93.8	473	22	AAAG64475	A dimeric anti-CD2
32	1251	93.8	475	18	AAW11641	Human anti-RSV mon
33	1251	93.8	476	20	AAW88464	Monoclonal antibod
34	1250	93.7	399	21	AAAY70867	Human interferon-b
35	1250	93.7	423	21	AAAY70869	Human interferon-b
36	1250	93.7	446	15	AAAY58753	VCAM 2D-IgG. Homo
37	1250	93.7	446	20	AAAY23986	VCAM 2D-IgG. a sol
38	1250	93.7	446	20	AAAY01037	VCAM 2D-IgG protei
39	1250	93.7	446	20	AAW96743	A VCAM 2D-IgG1 fus
40	1250	93.7	473	22	AAAG64469	Human type antihum
41	1250	93.7	473	22	AAAG64471	Human type antihum
42	1250	93.7	476	14	AAAR31023	Antibody D heavy c
43	1250	93.7	481	13	AAAR24442	Sequence of antibo
44	1250	93.7	592	22	AAAB83838	Amino acid sequenc
45	1250	93.7	963	19	AAW70540	Integrin beta-1 ch

ALIGNMENTS

RESULT 1

AAW70539

ID AAW70539 standard; Protein; 1218 AA.

XX

AC AAW70539;

XX

DT 26-JAN-1999 (first entry)

XX

DE Integrin alpha-4 chain.

XX

KW Integrin; alpha-4 chain; immunoglobulin; chimeric; heterodimer complex;

KW inhibitor; binding; ligand; blood platelet; hemostatic; diagnostic agent;

KW human.

XX

OS Homo sapiens.

XX

FH Key

FT Peptide

FT Protein

FT

XX

PN WO9832771-A1.

XX

PD 30-JUL-1998.

XX

PF 29-JAN-1998; 98WO-JP00370.

XX

PR 29-AUG-1997; 97JP-0234544.

PR 29-JAN-1997; 97JP-0015118.

XX

PA (TORA ) TORAY IND INC.

XX

Location/Qualifiers  
1...39  
/note= "signal peptide"  
40..1218  
/note= "mature protein"



Qy	261	-----	260
Db	23978	SAWKSNERIKDKQFTIGGLEATEYEFVFAENETGLSRPRRTAMSIKTKLTSGEAPG	24037
Qy	261	-----	260
Db	24038	IRKEMKDVTTKLGEAQLSCQIVGRPLDIKWYRFGKELIQSRKYKMSDGRTHLTVMT	24097
Qy	261	-----	260
Db	24098	BEQDEGVYTCIATNEVGEVETSSKLLQATPQFHPGYPPLKERYYGAVGSTLRHLHWYIG	24157
Qy	261	-----	260
Db	24158	RPVPMTWPHGQKLLQNSENITIENTEHTYHLVMKNVQRKTHAGYKVQLSNVFGIVDAI	24217
Qy	261	-----	260
Db	24218	LDVEIQDKPKDPTGPIVIEALLKNSAVISWKPADDGGSWITNYVVEKCEKEGAEWQLV	24277
Qy	261	-----	260
Db	24278	SSAISVYTCRIVNLNENAGYFRSAQNTFGISDPLEVSSVWIIKSPFEKPGAPGKPTIT	24337
Qy	261	-----	260
Db	24338	AVTKDCVWAKPPASDGGAKIRNYYLEKREKKQNKWISVTTEEIRETVFSVKNLIEGLE	24397*
Qy	261	-----	260
Db	24398	YEFVKCENLGSEWSEISEPTPKSDVPLOAPHEKEELRNLRVYQSNATLVCKVTGH	24457
Qy	261	-----	260
Db	24458	PKFIVKWYRQKEIIADGLKYRQEFKGYHQLIIASVTDATVYQVRATNOGGSVSGT	24517
Qy	261	-----	260
Db	24518	ASLEVEVPAKIHLPKLTLEGMAVHALRGEVWSIKIPFGKPDPTITWQKGQDLIDNNGHY	24577
Qy	261	-----	260
Db	24578	QVIVTRSFSLVPNGVERKDAGFYVVCARNRFGIDQKTVELDVADVPDPPRGVKVSDVS	24637
Qy	261	-----	260
Db	24638	RDSVNLWTPEPASDGGSKITNYIVEKCATAERLVRGQARETRYVINLFGKTSYQFRV	24697
Qy	261	-----	260
Db	24698	IAENKFGLSKPSEPPTITTKEDKTRAMNYDEEVDRETREVSMTKASHSSTKELYEKMIA	24757
Qy	261	-----	260
Db	24758	EDLGRGEFIVHRCVETSSKTTYMAKFVKVGDQVLVKKEISILNIARHRNILLHESF	24817
Qy	261	-----	260
Db	24818	ESMEELVMIFEFGSLDIFERINTSAFELNEREIVSVHQVCEALQFLHSHNIGHFDIRP	24877
Qy	261	-----	260
Db	24878	ENIIYQTRRSSTIKIEFGQARQLKPGDNFRLLFTAPEYYAPEVHQHVDVSTATDMWSLG	24937
Qy	261	-----KEAF	264
Db	24938	TLVYVLLSGINPFLAETNQOIENIMNAEYTFDEEAF	24974

Db 21758 GPPTVVKVTDSTKTTVSLWSPKVPDGGMEIIGYIEMCKADLDGWHKVNACVKTRYT 21817  
QY 248 ----- 247  
Db 21818 VTDLQAGEYKFRVSAINGACKGDSCEVTGTIKAVDRLTAPELDIDANFKQTHVVVRAGAS 21877  
QY 248 ----- 247  
Db 21878 IRLFIAYOGRPTPTAVWSKPSNLSLRADIIHTDSFSFLTIVENCNRNDAGKYTLTVENNS 21937  
QY 248 ----- 247  
Db 21938 GSKSITFTVKVLDTPGPPGPITFKDVTGRGSATLMDAPLLDGGARIHHYVVEKREASRRS 21997  
QY 248 ----- 247  
Db 21998 QWVISEKCTRQIFKVNDAEAGVPYIFRVSAVNEYGVPYEMPEFIVATEQPAPPRLDV 22057  
QY 248 ----- 247  
Db 22058 VDTSSSAVLAWLKPDDHGGSRITCYLLEMRQKGSDFWVEAGHTKQLFTTVERLVEKTEY 22117  
QY 248 ----- 247  
Db 22118 EFRVAKNDAGYSEPREAFSSVIIKEPOIPTADLTGITNQLITCKAGSPFTIDVPISGR 22177  
QY 248 ----- 247  
Db 22178 PAPKVTWLEMRUKETDRVSITTTKDRTTLTIVKDSMRGDSGRYFLTLENTAGVKTFEVS 22237  
QY 248 ----- 247  
Db 22238 VVIGRCPVTPGPIEVSSVSAESCVLSWGPCKDGGTEITNYIVEKRESGTTAMOLVNS 22297  
QY 248 ----- 247  
Db 22298 VKRQIKVTHLTKYMEYFRVSSNRFVSKPLESAPIIAEHPPVPPSAPTRPEVYHSA 22357  
QY 248 ----- 247  
Db 22358 NAMSIRWEEPYHDGGSKIIGYWEKERNILWKENKVPCLCNYKVTGLVEGLEYOFR 22417  
QY 248 ----- 247  
Db 22418 TYALNAGVSKASEASRPIAQNPDAPGREVTDVTRSTVSLIWSAPAYDGGSKVVGVI 22477  
QY 248 ----- 247  
Db 22478 IERKPVSEVGDGRWLKCNYYTIVSDNFTVTALSEGDTYEFRLAKNAAGVISKSESTGP 22537  
QY 248 ----- 247  
Db 22538 VTCRDEYAPPKAELDARLHGLVITIRAGDLVLDAAVGGKPEPKIITWTKGDKELDCEK 22597  
QY 248 ----- 247  
Db 22598 SLOVTKRATAVIFCDRSDSGKYLTVKNASGTKAVSMVKVLDSPGCGKLTFSRVTO 22657  
QY 248 ----- 247  
Db 22658 EKCTLAWSLPQEDGAEITHYIVERRETSLRNWVIVEGECTLSYVVTRLIKNNYIFRV 22717  
QY 248 ----- 247  
Db 22718 RAVNKYGPVPESEPIVARNSETIPSPGIPPEVGTGKEHIIQWTKPESDGGNEISNY 22777  
QY 248 ----- 247  
Db 22778 LVDKREKSLRWTRVNDYVYDTRLVKTVSLMEGCDYQPRVTAVNAAGNSEPSEASNFIS 22837  
QY 248 ----- 254  
Db 22838 CREPSYTPGPPSAPRVVDTTKHSISLAWTKPMYDGGTDIVGYVLEMQEKDIDQWYRVHTN 22897  
WLKAFYD  
| | |

QY 255 ----- 254  
Db 22898 ATIRNTEFTVPLDKMGQKYSFRVAANVKGMEYSESIAETIEPVERIEIPDELADLLK 22957  
QY 255 ----- 254  
Db 22958 TVTIRAGASLRLMVSVSGRPPPVITWSKOGIDLASRAIIDTTESYSLIIVDKVNRDAGK 23017  
QY 255 ----- 254  
Db 23018 YTIEAENSGKKSATVLVKVYDTPGCPSPKVKVSRDSVTITWEIPTIDGAPVNNIV 23077  
QY 255 ----- 254  
Db 23078 EKREAAAMRAFKVTTKCKSKTLYRISGLVEGTWYFRVLPENIYGIGEPCESTDVAVLVSEV 23137  
QY 255 ----- 254  
Db 23138 PLVPAKLEVVDVTKSTVTLAMEKPLYDGGSRLTGYVLEACKAGTERMKVVTTLKPTVLEH 23197  
QY 255 ----- 254  
Db 23198 TVTSLNEGEQYLFRIARAONEKGVSEPRETVTAVTVQDLRVLPTIDLSTMPOKTTHVPAGR 23257  
QY 255 ----- 254  
Db 23258 PVELVPIAGRPAPPAASWFFAGSKLRESERVTVETHTKVAKLTIRETITRTDGETLLEK 23317  
QY 255 ----- 254  
Db 23318 NVTGTTSETIKVIIIDKPGPTGPIKIDEIDATISITISWEPPELGGAPLSGYVVEQORDA 23377  
QY 255 ----- 254  
Db 23378 HRCGLVPSVSRSTFKFTRLTEGNEYVFRVAATNRFSGYLSQSEVIECRSSIRIPGP 23437  
QY 255 ----- 254  
Db 23438 PETLIQFDSRDGMLTWYPEDDGGSOVTGYIVERKEVRADRVVRVKNKVPVTMYRST 23497  
QY 255 ----- 254  
Db 23498 GLTEGLEYHRVTAINARGSKPSRSPKPIVAMDPIAPPGKPNQPRVTDTRTSVSLAWS 23557  
QY 255 ----- 254  
Db 23558 VPEDEGSKVTGYLLEMQVDQHEWTKCNTPTPKIREYTLTHLPQGAERFVLACNAGG 23617  
QY 255 ----- 260  
Db 23618 PGEPAEVPGTVKVTMELEYDPYELDERYQEGIFVROGGVIRLTIPIKGPFPICKWTKEG 23677  
QY 261 ----- 260  
Db 23678 QDISKRAMIATSETHTELVIKEADRGDSGTVDLVLENKCGKAVYIKRVVIGSPNSPEGP 23737  
QY 261 ----- 260  
Db 23738 LEYDDIQVRSVRVSWRPPADGGADILGYILERREVPKAAWYTIIDSRVGTSLVVGKLE 23797  
QY 261 ----- 260  
Db 23798 NVEYHFRVSAENQFISKPLKSEEPVTPKTLNPPEPPSNPPEVLDVTKSSVLSWSRPK 23857  
QY 261 ----- 260  
Db 23858 DGGSRVTGYIERKETSTDKVVRHNKTQITTTMYTGLVPOAEXQFRIIAQNDVGLSE 23917  
QY 261 ----- 260  
Db 23918 TSPASEPVVCKDPDFKPSQPGELEILISKDSVTLQWKEPCDGGKEILGYWVEYRQSGD 23977

Qy	245	-----	244
Db	19598	SGKTTITIAWKPFDGGAPITGYVEYKKSDDTDWKTQSIOSLRCTEYITISGLTGAEVV	19657
Qy	245	-----	244
Db	19658	FRVSVNVKVASDPSDSDPOIAKEREEPLFDIDSEMRKTLIVKAGASFTMTVPFRGRP	19717
Qy	245	-----	244
Db	19718	VPNVLSKPDOTDLTRAYVDTDSRTSLTIENANRNDGKYTLTIQNVLSAASLVLVVKV	19777
Qy	245	-----	244
Db	19778	LDTGPPPTNITVDYTKESAVLSWDVPENDGGAPVKNYHIEKREASKAWSVTNNCNRL	19837
Qy	245	-----	244
Db	19838	SYKVTNLQEGAIYFVRVSGENEFVGVI PAETKEGVKITEKPSPEKLGVTISIKDSVSLT	19897
Qy	245	-----	244
Db	19898	WLKPEHGGSRIVHYVVALEKQGNWVKCAVAKSTHRVVSGLRENSEYFFRVFAENQAG	19957
Qy	245	-----	244
Db	19958	LSDPRELLPVLLIKEQLEPPIDMKNFPSHTVYVRAGSNLKVDPISGKPLPKVTLSDRG	20017
Qy	245	-----	244
Db	20018	VPLKATRFNTEITAENLTINLKESVTADAGRYEITANSSTTKAFINIVVLDROPGPPT	20077
Qy	245	-----	244
Db	20078	GPVVISDITESVTLKWPPEKYDGGSVQNTYILLKRETSTAVTEVSATVARTMMKVMKL	20137
Qy	245	-----	244
Db	20138	TTGEEYQFRIKAENRFGISDHDSACVTVKLPYTPPGPSTPWNTVNTRESITVGWHEPV	20197
Qy	245	-----	244
Db	20198	SNGSAVVGYHLEMKDRNSILWQANKLVIRTHFKVTTISAGLIYEFVRYAENAAGVGK	20257
Qy	245	-----AFD-----	247
Db	20258	PSHPSEPVLADACEPPRNRVITDISKNSVLSWQQPAFDGGSKITGYIVERRDLPDGRW	20317
Qy	248	-----	247
Db	20318	TKASFTNVTETQFIISGLTQNSQYEFVRFARNAVGSIISNPSEVVGPTICIDSYGPPIDL	20377
Qy	248	-----	247
Db	20378	PLEYEVVYKRGTSVKLRAGISGKPAPTIEWYKDDKELQNALVCVENTTDLASILIKD	20437
Qy	248	-----	247
Db	20438	ADRLNSGCEYELKLRNMGASATIRVQILDKPPGPPGPIEFTVTAEKITLLWRPPADDG	20497
Qy	248	-----	247
Db	20498	GAKITHYIVEKRETSRVVWSMVSEHLEBIIITTKIIGNEYIFRVRVKNYGIGEPLES	20557
Qy	248	-----	247
Db	20558	DSVAKNAFVTPGPGPIEVTKITKNSMTVVWSRPIADGGSDISGYFLEKRDKSLGWFK	20617
Qy	248	-----	247
Db	20618	VLKETIRDTROKVTGLTENSQYQYRVCVNAAGOGPSEPSEFYKAADPIDPPGPPAKIR	20677
Qy	248	-----	247

Db	20678	IADSTKSSITLGWSKPVYDGGSAVTGYVVEIROGEEBETTVSTKGEVTRTEYVVVSNLKP	20737
Qy	248	-----	247
Db	20738	GVNTYFRVSAVNCAGQGEPIEMNEFPQAKDILEAPEIDLDVALRTSVIAKAGEDVQVLIP	20797
Qy	248	-----	247
Db	20798	FKGRPPPTVTRKDEKNLGSADARYSIENTDSSSLTTPQVTRNDTGKYLTIENGCVGEPK	20857
Qy	248	-----	247
Db	20858	SSTYSVKVLDTPAACQKLOVKHVSRGVTLLWDPPLIDGGSPIIINYVIEKRDATKRTWSV	20917
Qy	248	-----	247
Db	20918	VSHKCSSTSFXLIDLSEKTPFFFRVLAENEIGIGPCETTEPFVKAEEVPAPIRDLMSKDS	20977
Qy	248	-----	247
Db	20978	TKTSVILSWTKPDPDGGSVITEYVVERKKGGEQWTHAGISKTCEIEVSQLEKQSVLEFR	21037
Qy	248	-----	247
Db	21038	VFAKNEKGLSDPVTIGPITVKELIITPEVDLSDIPGAQVTVRIGHNVHLELPYKGPKP	21097
Qy	248	-----	247
Db	21098	ISWLKDGILPKESEFVRFSKTENKITLSIKNAKHEGKGYTVILDNAVCRVIAVITVITL	21157
Qy	248	-----	247
Db	21158	GPPSKPGPIRFDEIKADSVILSDVPEDNCGGEITCYISIEKRETSQTNWKNWCVSVART	21217
Qy	248	-----	247
Db	21218	TFKVPNLVKAQYQFVRAENRYGVSQPLVSSIIIVAKHQFRIIPGPGKPIYVNTSDGMS	21277
Qy	248	-----	247
Db	21278	LTWDAPVYDGGSEVTGFHVEKKERNILWQKVNTSPISGREYRATGLVEGLDYQFRVAAE	21337
Qy	248	-----	247
Db	21338	NSAGLSSPSPDSKFTLAVSPVDPPTDYIDVTRTITLKNPPLRDGGSKIVGYSIEKR	21397
Qy	248	-----	247
Db	21398	QGNERNVRCNFTDVSECQYTVTGLSPGDRYEFRIIARNAVCTISPPSQSSGIIWTRDENV	21457
Qy	248	-----	247
Db	21458	PPIVEFGPEYDGLIIKSGESLRIKALVQGRPVPRVTFWFKDGBIEKRMNMEITDVLGST	21517
Qy	248	-----	247
Db	21518	SLFVRDATRDRHGVYTVVEAKNWSAKAEIKVKVQDTPGKVGPIRFTNITGKMTLWWD	21577
Qy	248	-----	247
Db	21578	APLNDGCAPITHYIIEKRETSRLAWALIEDKCEAQSVTAIKLINGNEYQFRVSAVNRFGV	21637
Qy	248	-----	247
Db	21638	GRPLDSDPVVAQIOYTVDPAGPIPEPSNITGNSITLTWARPESDGGSEIOQYILERREKK	21697
Qy	248	-----	247
Db	21698	STRVVKVISKRPISETRFKVTGLTEGNEYEFHVMAENNAAGVGPASGISRLIKREPVNPP	21757
Qy	248	-----	247



Db	17378	DLRLVTLGLTEDHEYFRVSAENAGVGEPSATVYKACDPVFKPGPPTNAHIVDTTKN	17437
QY	241	-----	240
Db	17438	SITLANGKPIYDGGSEILGVVVEICKADEEMQIVTPQGLRVTRFEISKLTIEHQYKIR	17497
QY	241	-----	240
Db	17498	VCALNKVGLGEATSVPGTVKPEDKLEAPDLDELSELKGIIVRAGGSARIHIPFKGRPTP	17557
QY	241	-----	240
Db	17558	EITWSREGEFTKVOIEKGVNTQLSIDNCDRNDAGKYLKLENSSGSKSAFVTVKVLID	17617
QY	241	-----	240
Db	17618	TPGPPQNLAVKEVKDSAFVWEPIIDGGAKVKNVIDKRESTRKAYANVSSKCSKTSF	17677
QY	241	-----	240
Db	17678	KVENLFEAIYPRVMAENEFVGVVETVDVAKAAEPPSPGKVTLTDOVSQTSASLMWE	17737
QY	241	-----	240
Db	17738	KPEHDGSRVLGVVEMQPKTEKWSIVAESKVCNAVVTGLSSGOEYQFVRKAYNEKGKS	17797
QY	241	-----	240
Db	17798	DPRVLGVPVIAKDLTIQPSLKLFPNTYSIQAGEDLKEIPVIGRPRPNISWVKDGEPLAQ	17857
QY	241	-----	240
Db	17858	TTRVNVEETATSVLHIKEGNDKDFGKYVTATNSAGTATENLSVIVLEKPGPVGPVRF	17917
QY	241	-----	240
Db	17918	DEVSAFVVISWEPPAYTGCQISNIVIEKRDTTTTNHWVSATVARTIKITKLGTGE	17977
QY	241	-----	240
Db	17978	YQPRIFAENRYKSAFLDSKAVIVQYPFKEPGPGPFVTSISKQMLVQWHEPVNDGOT	18037
QY	241	-----	240
Db	18038	KIIGYHLEQEKNSILWVKNLKTPIODTRFKTGTGLDEGLEEFKVSANIVIGIKPSKVS	18097
QY	241	-----	240
Db	18098	ECFVARDPCDPPORPEAIVITRNNVTLKWKPAYDGGSKITGYIVEKKDLDPGRWMAKF	18157
QY	241	-----	240
Db	18158	TNVLETEFTVSLGLVEDQRYEVRVIARNAAGNFSEPSDSSGATARDEIDAPNASLDPKYK	18217
QY	241	-----	240
Db	18218	DVJVVHAGETFVLEADIRKPIPDVVVWSKDGELEETAARMEIKSTIQKTLVVKDCIRT	18277
QY	241	-----	240
Db	18278	DGGQYILKLSNVGGTKSIPITVKVLDROPPEGPKLVGTGTAECYLAWNPPLODGGANI	18337
QY	241	-----	240
Db	18338	SHVIEKRETSRLSWTQVSTEQALNYKVTKLLPGNEYIFRVMVKNKYGIGEPLESGPYT	18397
QY	241	-----	240
Db	18398	ACNPYKPPGPPSTPEYSAITKDSMVVTWARPVDDGGTEIEGYILEKRDKEGVRTKCNKK	18457
QY	241	-----	240
Db	18458	TLTDLRLVTLGLEHSHYEFVRVAENAGVGEPSFVYFACDALYPPGPPSPNPKVTDT	18517

QY	241	-----	KLKE-----	244
Db	18518	SRSSVSLANSKPIYDGGAPVKGYVVEVEKAAADEWTTCTPTGLQCKQKQTVVTKLKENY	18577	
QY	245	-----	1111	244
Db	18578	NFRICAINSEGVGEPATLPGSVVAQERIBPPEIELDADLRKVVVLRASATLRLFVTIKGR	18637	
QY	245	-----	-----	244
Db	18638	PEPEVKWEKAEGLTDRAGIEVTSSTMLVLDNVTRFDSGRYNLTLENSSGSKTAFVNR	18697	
QY	245	-----	-----	244
Db	18698	VLDSPSPVNLTIREVVKDSVTLWEPLIDGGAKITNYIVEKRETRKAYATITNCTK	18757	
QY	245	-----	-----	244
Db	18758	TTFRIENLQGGCSYFPRVLASNEYIGLPAETTEPVKVSEPLPPGGRVTLVDVTRNTATI	18817	
QY	245	-----	-----	244
Db	18818	KWEKPESDGGSKITGYVVEVEMOTKGSEKWSCTQVKTLEATISGLTAGEEYVFRVAANEK	18877	
QY	245	-----	-----	244
Db	18878	GRSDPRQLGVPVIARDIEIKPSVELPFHTFNVKAREQLKIDVPFKGRQATVNRKDGOT	18937	
QY	245	-----	-----	244
Db	18938	LKETTRVNVSSKTVTSLSIKEASKEDVGTIELCVSNSAGSITPITIIVLDRPPGPI	18997	
QY	245	-----	-----	244
Db	18998	RIDEVSCDSITISWNPPEYDGGCQISNYIVEKETTSTTWHIVSQAVARTSIKIVRLTTC	19057	
QY	245	-----	-----	244
Db	19058	SEVQFRVCAENRYKSSYSESSAVVAEYFPSPGPGTPKVVHATKSTMLVTWQVPVNDG	19117	
QY	245	-----	-----	244
Db	19118	GSRVIGYHLEYKERSILWSKANKILIAOTQMKVSGLDEGLMYERYVAENIAGIKCSK	19177	
QY	245	-----	-----	244
Db	19178	SCEPVAPDPCDPPGOPEVTNITRKSLSKWSKPHYDGGAKITGYIVERRELDPGRWLKC	19237	
QY	245	-----	-----	244
Db	19238	NYTNIQETYFEVTELTEDQRYEFRVFARNAADSVSEPSSTGPIIVKDDVPEPRVMDVK	19297	
QY	245	-----	-----	244
Db	19298	FRDVIIVKAGEVLKINADIAGRPLPVISWAKDGIEIERARTEIISTDNHTLLTVKDCIR	19357	
QY	245	-----	-----	244
Db	19358	RTGQVVLTKNVAGTRSVAVNCKVLDKPGPPAGPLEINGLTAEKCSLSWGRPOEDGGAD	19417	
QY	245	-----	-----	244
Db	19418	IDYIIVEKRETSHLAWTICEGELQMTSCKVTLLKKGNEYIFRVTVGNKYGVGEPLESVAI	19477	
QY	245	-----	-----	244
Db	19478	KALDPFTVPSPPTSLEITSVTRESMTCWSRPESDGGSEISGYIIEERKNSLRWRVKNK	19537	
QY	245	-----	-----	244
Db	19538	KPVYDLRVKSTGLREGCEYERYVAENAGLSLPSSETSLIRAEDPVFLPSPSPKPIVD	19597	

QY	226	-----	225
Db	15218	KVSLTEGHEYEFRIMAENAGISAPSPSPFYKACDTVFKPGPGNPRVLDTSRSSISI	15277
QY	226	-----	225
Db	15278	AWNKPYYDGGSEITGYMVEIALPEDEWQIVTPPAGLKATSYTITGLTENQEKIRIYAM	15337
QY	226	-----	225
Db	15338	NSEGLGEPALVPGTPKABDRMLPEIELDADLRKVVTIRACCTLRLEVPKIGRPAPEVKW	15397
QY	226	-----	225
Db	15398	ARDGESLDKASIBSTSSYILLIIVGNVNRFDGKYILTVENSNGSKSAFVNVRLDTPGP	15457
QY	226	-----	225
Db	15458	PODLKRVETKTSVTLTWDPPLLDGSKIKNYIVEKRESTRKAYSTVATNCHKTSMKVDQ	15517
QY	226	-----PGK-----	228
Db	15518	LOEGCSYFRVLAENEYIGLPAETAESVKASERPLPPGKITLMDVTRNSVLSWEKPEH	15577
QY	229	-----	228
Db	15578	DGGSRIILGYIVEMQTKGDKWATCATVKVTEATITGLIQEEYSFRVSAONEKGISDPQ	15637
QY	229	-----	228
Db	15638	LSVPVIAKDLVPPAPKLLFNTFTVLAGEDLKVDVPFGRPTPAVTWHKDNVPLKQTRV	15697
QY	229	-----	228
Db	15698	NAESTENNSLLTIKDACREDVGHVYVKLTNSAGEAIEFLANVILDKPGPPTGPKMDEV	15757
QY	229	-----	228
Db	15758	ADSTLSWGPYPKYDGGSSINNYIVEKRDTSITTWQIVSAIVARTTIKACRLKTGCEYQFR	15817
QY	229	-----	228
Db	15818	IAENRYGKSTYLNSEPTVAQYFPKVPGPPTPVVTLSSRDSMEVOWNEPISDGSRVIG	15877
QY	229	-----	228
Db	15878	YHLERKERNILWVKLNKTPIPOTKFTTGLEGEVEYFRVSAENIVGIGKPSKVSCEYV	15937
QY	229	-----	228
Db	15938	AROPCDPPRPEALIVTRNSVTLOWKKPTYDGGSKITGYIVEKKELPEGRWMKASFTNII	15997
QY	229	-----	228
Db	15998	DTHFEVTGLVEDHRYEPRVIARNAAGVFSEPSSESTGAIARDEVDPRIISMDPKYKDTIV	16057
QY	229	-----	228
Db	16058	VHAGESFKVDADYGRPIPTIOWIKGQELSNARLEIKSTDPAFSLSVKDAVRVDSGNY	16117
QY	229	-----	228
Db	16118	ILKAKNVAGERSVTNVKVLDRPGPPBPVVISGVTAEKCTLAWKPPLODGGSDIINYIV	16177
QY	229	-----	228
Db	16178	ERRETSRLVTVVDANVQTLSCRVTKLLEGNEYTFRIMAVNKYGVGEPLESEPVAKNPF	16237
QY	229	-----	228
Db	16238	VVPDAPKAPETVTTKDSMTVWVERPASDGGSEILGVVLEKROKEGIRWTRCHKRLIGEL	16297
QY	229	-----	228

Db	16298	RLRVTLGLIENHDYEFVSAENAGLSPPSPSAYOKACDPIYKPGPNPKVIDITRSSV	16357
QY	229	--DWLKAIFYD-----	236
Db	16358	FLSWSKPIYDGGCEIOGYIVEKCDVSYGWETMCTPPTGINKTNIEVEKLEKHEYNFRIC	16417
QY	237	-----	236
Db	16418	AINKAGVGEHADVPPTIIVEEKLAPDIDLDELURKIINIRAGSRLFLFVPIKGRPTPEV	16477
QY	237	-----	236
Db	16478	KWKGVDEIRDAALIDVTSSTSLVLDNVNRYDSGKYTLTLENSSGTKSAFVTVRVLDTDP	16537
QY	237	-----	236
Db	16538	SPPVNLKVEITKDSVSITWEPPLLDGGSKIKNYIVEKREATRKSAAVVVTCNCHKNSKI	16597
QY	237	-----KVAE-----	240
Db	16598	DLOEGCSYFRVTAENEYIGLPAQTADPIKVAEVPQPPGKITVDDVTRNSVLSWTKP	16657
QY	241	-----	240
Db	16658	EHDGSKIIQYIVEMQAKHSEKARVKSLOAVITNLTOGBEYLFRVVAVNEKGRSDP	16717
QY	241	-----	240
Db	16718	RSLAVPIAKDLVIEPDVKPAFSSYSVQGDCLKIEVPIISGRPKPTITWTKGDLPLKQTT	16777
QY	241	-----	240
Db	16778	RINVTSDLTLTUSIKETHKDDGGQYGITVANVVGOKTASIEIVTLDKPDPGPKVFKDD	16837
QY	241	-----	240
Db	16838	VSAESITLSWNPPLYTGGCQITNYIVQKRDTTITVMDVVSATVARTLKVTCLKTGTEYQ	16897
QY	241	-----	240
Db	16898	FRIFAENRYGQSFALSDPIVAQYYPYKEPGPGTPTAFATISKDSMVIOHHEPVNNGSPV	16957
QY	241	-----	240
Db	16958	IGYHLERKERNILWTKVKNKTIITHDTQFKAQNLBEGIEYEFVYAENIVGVGKASKNSEC	17017
QY	241	-----	240
Db	17018	YVARDPCDPPGTPEPIMVKRNEITLQWTKPVYDGGSMITGYIVEKRDLPDGRWMKASFTN	17077
QY	241	-----	240
Db	17078	VIETQFTVSGLTQDQYEFVRIAANAAGAIKSPSDSTGPIITAKDEVELPRISMDPKFRDT	17137
QY	241	-----	240
Db	17138	IVVNAGETFRLEADVHGKPLPTIENLRGDKIEESARCEIKNTDFKALLIVKDAIRIDGG	17197
QY	241	-----	240
Db	17198	QYILRASNVAGSKSPVNVKVLDRPGPEGPVQVGTVTSKCSLTWSPPLQDGGSDISHY	17257
QY	241	-----	240
Db	17258	VVEKRETSRLAWTVVASEVVVNTSLKVTKLLEGNEYVFRIMAVNKYGVGEPLESAPVLMKN	17317
QY	241	-----	240
Db	17318	PFVLPGPSPKSVLENTIAKDSMTVCWNRPDSDGGSEILGIYIVEKDRSGIRWIKCNKRIT	17377
QY	241	-----	240

Db 12998 VISNITKDHMTVSWKPPADDDGGSPITGYLLEKRETOAVNMTKVNKRKPIIERTLKATGLQE 13057  
QY 209 ----- 208  
Db 13058 GTEYEFRTAINKAGPKPSDASKAAYARDPOYPAPPAPFKVYDTRSSVLSWGKPAY 13117  
QY 209 ----- 208  
Db 13118 DGGSPILIGLVEVKRADSDNWRCNLQNLQKTRFETGLMEDTOYQFRVAVNKGISYD 13177  
QY 209 ----- 208  
Db 13178 PSDVPDKHYPKDLIPPEGELDADLRKLLILRAGVTMRLYVPVKGRPPKITTWSKPNVNL 13237  
QY 209 ----- 208  
Db 13238 RDRIGLDIKSTDFRFLRCENVNKYDAGKYILTLENSCGKKEYTIWVKYLDTPGPPVNT 13297  
QY 209 ----- 208  
Db 13298 VKEISKOSAYVTWEPPIIDGGSPIIYVQKRAERKSWSTVTTECSKTSFRVANLEBK 13357  
QY 209 ----- 208  
Db 13358 SYFFRFAENEXYIGDPCGTRDAVKASQTPGPVVDLKVRSVSKSSCSIGWKKPHSDGGSR 13417  
QY 209 ----- 208  
Db 13418 IIGYVDFTEENKQWVKMSLSLOYSAKDLTEGREYTFRVSANENGEGTPEITVVAR 13477  
QY 209 ----- 208  
Db 13478 DDVAPDLDLGLDPLCYLAKENSFRKIPKGPAPSVSWKKGEDPLATDTRVSVSS 13537  
QY 209 ----- 208  
Db 13538 AVNTTLIVYDCQSDAGYITLKNVAGTKEGTISIKVVGKPGIPTGPIKDFDEVTAEMT 13597  
QY 209 -----MHEALH----- 214  
Db 13598 LKWAPPKDDGGSEITNYILEKRDVNNKWTCASAVOKTTFRTVRLHGMMEYTFRVSABN 13657  
QY 215 ----- 214  
Db 13658 KYGVGEGLKSEPIVARHPFDPDAPPPNIVDVHRDSVSLTWTDPKKTGGSPITGYHLEF 13717  
QY 215 ----- 214  
Db 13718 KERNSLLMKRANKTPIRMDFKVTGLTGLYEYEFVRMAINLAGVGKPSLPSEPVVALDPI 13777  
QY 215 ----- 214  
Db 13778 DPPGKPEVINTRNSVTLIWTBPKYDGGHKLTYIVEKRDLPKSWMKANHVNPCEAPT 13837  
QY 215 ----- 214  
Db 13838 VTDLVEGKYEPRIAKNTAGIASPSESTETIICKDEYEAPTIVLDPTIKDGLTIKAGD 13897  
QY 215 ----- 214  
Db 13898 TIVLNAISTLGLKPLKSSWSKAGDIRPSDITQITSTPTSSMLTIKYATRKDAGEYTTA 13957  
QY 215 ----- 214  
Db 13958 TNPFGTKVEHKVTVLDVPGPGPVEISNVSAEKATLTWTPPLEDGGSPIKSYILEKRET 14017  
QY 215 -----NHY----- 217  
Db 14018 SRLMTVVSIEDIQSCRHVATKLIQNEXYIFRVSANVHYKGPVQSEPKVMVDREPGPGP 14077  
QY 218 ----- 217  
Db 14078 PEKPEVSNVTKNTATVSNKRPVDDGSGEITGYHVERREKSLRWVRAIKTPVSDLRCKVT 14137

QY 218 -----TOKSLSL----- 225  
Db 14138 GLQEGSTYEFVRYSAENRAGIGPPSEASDVLKMDAAYPGPPSPSHVHTDTTKKSASLAWG 14197  
QY 226 ----- 225  
Db 14198 KPHYDGLLEITGYVVEHQKVGDEAWIKDTGTALRITQFVVFDLOTQEKYFNFRISAINDA 14257  
QY 226 ----- 225  
Db 14258 GVGEPAVIDDVEIVEREMAPDFELDELRLTLVVRAGLSIRIFVPKGRPAPEVTTWKDN 14317  
QY 226 ----- 225  
Db 14318 INLKNRANIENTESFTLLIIPECNRYDTGKFVMTIENPAGKSGFVNVRLDTPGPVLNL 14377  
QY 226 ----- 225  
Db 14378 RPTDITKDSVTLHWDLPLIDGGSRTNYIVEKREATRKSYSTATTKCHKCTYKVTLGSEG 14437  
QY 226 ----- 225  
Db 14438 CEYFFRYMAENEYIGCEPTEPTEPVKASEAPSPDPSLNIMDITKSTVSLAWPKPHDGG 14497  
QY 226 ----- 225  
Db 14498 KITGVIEAQRKGSQDOWHITTVKGLECVVRNLTEGEYTFQVMVNSAGRSAPRESRPV 14557  
QY 226 ----- 225  
Db 14558 IVKEQTMPLPELDLRGIYQKLVIAKAGDNITKVEIPVLRPKPTVTVWKKGQILKQTORVNF 14617  
QY 226 ----- 225  
Db 14618 ETTATSTILNINECVRSDSGPYPLTARNIVGEVDVITIQVHDIPOGPTGPIKDFEVSSD 14677  
QY 226 ----- 225  
Db 14678 FVTFSDPPDPPDGGVPISNYVVMRODSTTWVELATTVIRTYKATRLTTGLQYQFRVK 14737  
QY 226 ----- 225  
Db 14738 AQNRVGVGPGITSACIVANYPFKVPGPPPTQVTAVKDSMTISWHEPLSDGSGSPILGYH 14797  
QY 226 ----- 225  
Db 14798 VERKERNGILWQTVSKALYVGNIFKSSGLTDGIAEYEFVRIAENMACKSKPSKSEPMLAL 14857  
QY 226 ----- 225  
Db 14858 DPIDPPGKPVPLNITRHTVTLKWAKPEYTGFGFKITSYIVEKRDLPNGRMLKANFSNILEN 14917  
QY 226 ----- 225  
Db 14918 EFTVSGLTEDAAEYEFVRIAENMACKSKPSKSEPMLAL 14977  
QY 226 ----- 225  
Db 14978 AGEAFRLEADVSGRPPPTMWSKDGKEGTAKLEIKIADFTNLVKNKDSRDSGAYTL 15037  
QY 226 ----- 225  
Db 15038 TATNPGGFAKHIFNVKVLDRPGPPEGLAVTEVTSEKCVLSWFPPLDDGGAKIDHIVOK 15097  
QY 226 ----- 225  
Db 15098 RETSLANTVNAVSEVOVTKLVTKLLKGNEYIFRVMVANKYGVGEPESEPLAVNPYGP 15157  
QY 226 ----- 225  
Db 15158 PDPPKNPEVTTITKDSMVVWGHGHPDSDGSEIINIVERRDKAGORWIKCNKTLTDLRY 15217

```
QY 177 ----- 176
Db 10838 DGGSVITNVERRDVASQWSPSATSKKSHFAKHLNEGQYLFRVAENQYGRGPF 10897
QY 177 ----- 176
Db 10898 VETPKPIKALDPLHPGPKDLHHVDVDKTEVSLVWKNKPDGSGSPITGYLVEYQEBGTQ 10957
QY 177 ----- 176
Db 10958 DWIKFKVTNLECVVTGLQOGKTYFRVKAENIVGLGLPDTTPIECQEKLVPPSELDV 11017
QY 177 ----- 176
Db 11018 KLIEGLVVKAGTVRFPAILIRGVPVPTAKWTTDGESEIKTDEHVTVDNFSSVLTKNCL 11077
QY 177 ----- 176
Db 11078 RRTGEYOITVSNAGSKTVAVHLTVLDVPGPPTGINILDVTPEHMTISWQPPKDDGGS 11137
QY 177 ----- 176
Db 11138 PVNIYVEKQDTRKDTGWVSSGSSKTKLIPHLOKCEYVFRVRAENKICVGPPLDSTP 11197
QY 177 ----- 176
Db 11198 TVAKHKFSPSPGPKPVVTDITENAAVTSWTLPKSDGSGSPITGYMERREVTKWVRVNK 11257
QY 177 ----- 176
Db 11258 TPIADLKFRVTGLYEGNTYFRVFAENLAGLSKPSSPDIKACRPKPPGPPINPKLD 11317
QY 177 ----- 176
Db 11318 KSRETADLVWTKPLSDGSGSPILGYVVEQCQPGTAQNRINKDELIROCAFRVPLIEGNE 11377
QY 177 ----- 176
Db 11378 YRFRIKAANIVGEGEPRELAESVIAKDILHPPEVELDVTCDVITVRVGOTIRILARVKG 11437
QY 177 ----- 176
Db 11438 RPEPDITWTKEGVLVREKRVLDIQLPRVELOIKEAVRADHGKYYIISAKNSSGHAQGSA 11497
QY 177 ----- 176
Db 11498 IVNVLDPRGPCQNLKVTNVTKENCTISWENPLDNGGSEITNFIVEYRKNQKGSIVASD 11557
QY 177 ----- 176
Db 11558 VTKRLIKANLANNEYFRVCAENKVGVGPTIETKTPILAINPIDRGPENLHIADKKG 11617
QY 177 ----- 176
Db 11618 TFVYLKRRPDYDGGSPNLSYHVERRLKGSDDWERVHKGSIKETHYMWDRVCVENQIYEF 11677
QY 177 ----- 180
Db 11678 VQTKNEGGESDWKTEEVVVKEDLQKPVLDLKLGSGLTVKAGDTIRILEAGVRGKPFPEVA 11737
QY 181 ----- 193
Db 11738 WTKDKDATDLTRSPRKIDTRADSSKFSLTAKRSDGKYYVVTATNTAGSFVAYATVNVL 11797
QY 194 ----- 193
Db 11798 DKPCPVRLNKIVDVSSDRCTVCWDPDEDGGEIQNYILEKETKRWVWSTYSATVLTGP 11857
QY 194 ----- 195
Db 11858 TTVTRLIEGNEYIFRVAENKIGTGPPTESKPVIAKTKYDKGRPDPPPEVTKVSKKEMTV 11917
QY 196 ----- 195
```

```
Db 11918 VWNPEYDGGKSGITGYFLEKKEKHSTRWVPVKNKSAIPERRMKVQNLPLDHEYOFRVKAEN 11977
QY 196 ----- 195
Db 11978 EIGIGEPSLSPRVVAKDPIEPPGPTNFRVDTTKHSITLGMGKPYVDGAPIIGYVVE 12037
QY 196 ----- 195
Db 12038 MRPKIADASPEGKRCNAAAQLVRKEFTVTSLDENQEYEFVCAQNOVGIGRPAELKEA 12097
QY 196 ----- 195
Db 12098 IKPKEILEPPEIDLDASMRKLIVIRAGCPIRLFAIVRGPRAPKVTWRKVGIDNVVRKGQV 12157
QY 196 ----- 195
Db 12158 DLVDTMAFLVIPNSTRDDSGKYSLLTLVNPAGKAVFVNVVRVLTDPGVSOLKVSVDVTKTS 12217
QY 196 ----- 195
Db 12218 CHVSWAPPENDGSGOVTHYIVKREADRKTWSTVTPVKKTSFHVNTLVPCNEYFPRVTA 12277
QY 196 ----- 195
Db 12278 VNEYGCVPTDVPKPVLASDPLSEPDPPRKLEVTMTKNSATLAWLPLRDGGAKIDGYI 12337
QY 196 ----- 195
Db 12338 TSYREEEQADRWTEYSVWKDLSLVVTGLKEGKKYFRVAARNAVGSVSLPREAGVYEAK 12397
QY 196 ----- 202
Db 12398 EQLLPPKILMPEOITIKAGKKLRIEBAHVYCKPHPTCKWKKAGEDEVTSSSHVLAHVHKADSSS 12457
QY 203 ----- 202
Db 12458 ILIIKDVTRKDSGYSLTAENSSGTDQKIKVVVMDAPGPPQPFDDISDIDADACSLSWH 12517
QY 203 ----- 202
Db 12518 IPLEDGGSNITNYIVEKCDVSRGDDWTALASVTKTSRVRGLIPGQYEIFRVAENRFGI 12577
QY 203 ----- 202
Db 12578 SEPLTSPKVAQPPFGVSEPKNARVTKNKDCIFVANDRPDSDGSGSPIIGYLIERKERN 12637
QY 203 ----- 202
Db 12638 SLLWVRANDFLVRSTEYPCAGLVEGLEYSFRIYALNKGSSPPSKPTEYVTRAMPVDPGP 12697
QY 203 ----- 202
Db 12698 KPEVIDVTKSTVSLIWARPKHDGSKIIGYFVEACKLPGDKWVRCNTAPHQIPQEEYAT 12757
QY 203 ----- 202
Db 12758 GLEEKAOYOFRAIARTAVNISPPSEPSDPVTILAENVPPIDLVAMKSLTLTVKAGTNVC 12817
QY 203 ----- 202
Db 12818 LDATVFGKPMPTVSWKKDGTLLKPAEGIKMAMQNLCTLELFSVNRKDSGDTITAENSS 12877
QY 203 ----- 202
Db 12878 GSKSATIKLVLDKPGPPASVKINKMYSDRAMLSWEPLEDGGSSEITNYIVDKRETSRPN 12937
QY 203 ----- 208
Db 12938 WAQVSATVITSCVSEKLEHGYQFRICAENKYGVGDPVTFEPATAKNPYDPPGRCDPP 12997
QY 209 ----- 208
```

```

RESULT 15
Q8WZB3
ID Q8WZB3 PRELIMINARY; PRT; 26926 AA.
AC Q8WZB3;
DT 01-MAR-2002 (TREMBLrel. 20, Created)
DT 01-MAR-2002 (TREMBLrel. 20, Last sequence update)
DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)
DE N2B-titin isoform.
GN TTN.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RN SEQUENCE FROM N.A.
RP MEDLINE=20309627; PubMed=10850961;
RX Freiburg A., Fromoltes K., Heil W., Cazorla O., Fougereousse F.,
RA Centner T., Kolmerer B., Witt C., Beckmann J.S., Gregorio C.C.,
RA Granzier H., Labeit S.;
RT "Series of exon-skipping events in the elastic spring region of titin
as the structural basis for myofibrillar elastic diversity.";
RL Circ. Res. 86:1114-1121(2000).
[2]
RN SEQUENCE FROM N.A.
RP MEDLINE=21573839; PubMed=11717165;
RX Bang M.L., Centner T., Fornoff F., Geach A.J., Gotthardt M.,
RA McNabb M., Witt C.C., Labeit D., Gregorio C.C., Granzier H.,
RA Labeit S.;
RT "The complete gene sequence of titin, expression of an unusual ~700
kDa titin isoform and its interaction with obscurin identify a novel
RT 2-line to 1-band linking system.";
RL Circ. Res. 89:1065-1072(2001).
DR EMBL: AJ277892; CAD12455.1;
DR InterPro: IPR000282; Cytok_receptor_2.
DR InterPro: IPR000719; Euk_pkinase.
DR InterPro: IPR000577; FGGY_kin.
DR InterPro: IPR003961; FN_III.
DR InterPro: IPR001092; HLH_basic.
DR InterPro: IPR003599; Ig.
DR InterPro: IPR003598; Ig_c2.
DR InterPro: IPR003600; Ig_lik.
DR InterPro: IPR003006; Ig_MHC.
DR InterPro: IPR002016; Peroxidase.
DR InterPro: IPR002290; Ser_thr_pkinase.
DR InterPro: IPR001245; Tyr_pkinase.
DR Pfam: PF00041; fn3; 132.
DR Pfam: PF00047; ig; 91.
DR Pfam: PF00069; pkinase; 1.
DR ProDom: PD000001; Euk_pkinase; 1.
DR SMART: SM00060; FN3; 132.
DR SMART: SM00409; IG; 111.
DR SMART: SM00409; IGc2; 93.
DR SMART: SM00410; IG_lik; 3.
DR SMART: SM00220; S_TKc; 1.
DR SMART: SM00219; TyrKc; 1.
DR PROSITE: PS00933; FGGY_KINASES_1; UNKNOWN_1.
DR PROSITE: PS00038; HELIX_LOOP_HELIX; UNKNOWN_1.
DR PROSITE: PS00290; IG_MHC; UNKNOWN_1.
DR PROSITE: PS00435; PEROXIDASE_1; UNKNOWN_1.
DR PROSITE: PS00011; PROTEIN_KINASE_DOM; 1.
DR PROSITE: PS00109; PROTEIN_KINASE_TYR; UNKNOWN_1.
SQ SEQUENCE 26926 AA; 2993285 MW; 169AB42637A7C1FB CRC64;

Query Match 34.9%; Score 466; DB 4; Length 26926;
Best Local Similarity 0.69; Pred. No. 0.00029;
Matches 155; Conservative 48; Mismatches 40; Indels 26150; Gaps 46;

QY 3 KTHTCpp----- 9
| | |
| | |
Db 98 KAETAPNFQRLQSMIVRGSGVRLQVRVTGIPTPVVKFYRDGAETQSLDFQISQEGD 157
QY 10 ----- 9

```

Db 1238 SESEAVESGFDLRKKNYRILEGMYTFHCKMSGYPLPKIAWYKDGKRRIKHGERYQMDFLQ 1297  
QY 34 ----- 33  
Db 1298 DGRASLRIPVLPEDRGYITAFASNIKGNACSGKLYVEPAAPLGAPTYIPTLEPVSRIR 1357  
QY 34 -----ISRTP----- 38  
: || :  
Db 1358 SLSPRSRSRPTRMSPARMSPARMSPARMSPARMSGRLEETDESOLERLYKPVFVLKP 1417  
QY 39 ----- 38  
Db 1418 VSFKLEGOTARFDLKVGRPMPTFWFHDGQOIVNDYTHKVYIKEDGTQSLIIVPATPS 1477  
QY 39 ----- 38  
Db 1478 DSGEWTVAQNRAGRSSISVILTVEAVEHQVPMFVEKLNKYNIKEGSQLEMKVRATGNP 1537  
QY 39 ----- 38  
Db 1538 NPDIVLKNSDIIVPHKPKIRIEGTGKAALKIDSTVSQDSAWYTATAINKAGRDTTRC 1597  
QY 39 ----- 38  
Db 1598 KVNVEVEFAPEPERKLIIPRGTYRAKETAAPELEPLHLRYGOEQWEGDLYDKEKQKRP 1657  
QY 39 ----- 38  
Db 1658 FFKKLTSLRLKRFPAHECERLTPIGDPMTVMVWNLHDGKPLEAANRLRMINEFGYCSLD 1717  
QY 39 -----EUTC----- 42  
: ||  
Db 1718 YGVAYSRDGIIITCRATNKYGTHTSATLIVKDEKSLVESQLPEGRKGLQRIEELERMA 1777  
QY 43 ----- 42  
Db 1778 HEGALTGVTDQEKQKQKPDIVLYPEPVRVLEGETARFCRVGTGPQPKVNYLNGQLIRK 1837  
QY 43 ----- 42  
Db 1838 SKRFRVRYDGIHYLDIVDCSYDTEGVKVTAEENPEGVIEHKVLEIQOEDFRSVLRAP 1897  
QY 43 ----- 42  
Db 1898 EPRPEFHVHPGKLOFEVQKQVDRPVDTTETKEWKLKRAERITHEKVPESBELSKFR 1957  
QY 43 ----- 42  
Db 1958 RTEEGYEAITAVELKSRKKDSYBELLRKTKDELLHMTKELTEBEKKALAEKGITIPT 2017  
QY 43 ----- 42  
Db 2018 FKPDKIELSPSEAPKIFERIOSQTVGQSDAHFRVVRVVGKDPPECEWYKNGVKIERSDR 2077  
QY 43 -----VVVDVSHED----- 51  
: || :  
Db 2078 IYWWPEDNVCVIRVDVTAEDSASIMVKAINIAGETSSHAPLLVOAKLIPTTQELQDV 2137  
QY 52 -----PEVKFNWYDGVVH----- 66  
: || :  
Db 2138 VAKEKDTMATFECETSEPPFVKYKWKDGMVEHGDYRMHSDRKVHFLSILITDTSDAED 2197  
QY 67 -----NAKT----- 70  
: ||  
Db 2198 YSCLVEDENVTAKLIVGAVVEFVKELQDIEVPSYSGELECIIVSPENIEGKWYND 2257  
QY 71 -----KPR----- 73  
: ||  
Db 2258 VELKNGKTYITSRGRONLTVKDVTKEDQGEYSPVIDGKKTCKLKMKPRPIALQGLS 2317  
QY 74 ----- 73  
Db 2318 DQKVCEDIVLEVKVSLESVEGVNMKGQEVQPSDRVHIIVIDKQSHMLLIEDMTKEDAG 2377

QY 74 ----- 73  
Db 2378 NYSFTIPALGLSTGRVSVYSDVITPLKDVNVIETRAVLECKVSPDVTSVKWLND 2437  
QY 74 ----- 73  
Db 2438 QIKPDDRVOAIVKGTQKORLVINRTHASDEGPYKLIIVGRVETNCNLSVEKIKIIRGLDIT 2497  
QY 74 ----- 73  
Db 2498 CTEONVVFEVLSHSGIDVLNFKDKEIKPSSKYKIEAHGKIYKLTVLNMMKDDEGKYT 2557  
QY 74 ----- 73  
Db 2558 FYAGENITSGKLTVAGGALSPLTDQTVAESQEAFFECEVANPDSKGEWLRDGHPLTN 2617  
QY 74 ----- 73  
Db 2618 NIRSSEGHKRRLIIAATKLDIGEYTYKVATSKTSAKLKVEAVKIKKTLKNLTVTETQD 2677  
QY 74 ----- 73  
Db 2678 AVFTVELTHPNVGVQWIKNGVLESNEKYAISVGTIYSLRIKNCIAVDESUYGFLGR 2737  
QY 74 ----- 73  
Db 2738 LGASARLHVEIVKIIKKPDVTALENATVAFESVSHDTPVVKWFHKNVEIKPSDKHLV 2797  
QY 74 ----- 73  
Db 2798 SERKVHKLMLQNIISPADAGEYAVVQLECKAKFVETLHITKTMKNIEVPETKTASPEC 2857  
QY 74 ----- 73  
Db 2858 EVSHPNVPSMLKNGVEIEMSEKFIIVOGKLUHOLIIMNTSTEDSAEYTFVCGNDQVSAT 2917  
QY 74 ----- 73  
Db 2918 LTVTPIMITSMLKDINAEEKDITTEFVTVNYEGISYKWLKNGVEIKSTDKCOMRTKLLTH 2977  
QY 74 ----- 73  
Db 2978 SLNIRNVHFGDAADYTFVAGKATSTATLYVEARHIEFRKHIDKIVLEKKRAMFECEVSE 3037  
QY 74 ----- 73  
Db 3038 PDITVQWKKDDQELQITDRIKIQEKYVHRLIIPSTRMSDAGKYTVVAGGNVSTAKLFE 3097  
QY 74 ----- 73  
Db 3098 GRDVRIRSIKKEVQVIEKQRAVVEFEVNEDDVDHMYKDGIEINFQVQERHKYVVERRIH 3157  
QY 74 ----- 73  
Db 3158 RMFISETRQSDAGEYTFVAGNRSSVTLYVNAPEPPVQLQELQPVTVQSGKPARFCAVIS 3217  
QY 74 -----EE----- 75  
: ||  
Db 3218 GRPQPKISWYKEEQLLSTGFKCKFLHDGQYETLLILIEAPPEDAAYVYTCENKNDYGVATTS 3277  
QY 76 ----- 75  
Db 3278 ASLSVEPEVVPDQEMPVYPPAIITPLQDITVTSQOPARFCQRCRVSGTDLKVSWSKDKK 3337  
QY 76 -----QYNSTYR----- 83  
: || :  
Db 3338 IKPSRFRMTQFEDTYQLEIAEAYPEDEGTFTFVASNAGVQVSSANLSLEAPESILHER 3397  
QY 84 ----- 83  
Db 3398 IEQEIEMEMKEFSSFLSABEEGLHSAELQLSKINETLELLESPPVYSTKFDSEKEGYGP 3457

QY 84 -----VSVL----- 87  
Db 3458 IFIKEYSNADISMGDVATLSVTVIGIPKPKIQWFFNGVLLTPSADYKRFVDPGDDHSLIL 3517  
QY 88 ----- 87  
Db 3518 FTKLEDEGEYTCMASNDYGTICSAYLKINSKGEGKHDTESESAAKSLKGLGPCPPHF 3577  
QY 88 -----TVLH----- 91  
Db 3578 LKELKPIRCAOGLPAIFEYTVWGPAPTIVTFWENKOLCTSVYVTHHNPGSGTFVND 3637  
QY 92 ----- 91  
Db 3638 POREDSGLYICKAENMLGESTCAAEALLVLEDTDMTPCKAKSTPEAPEDFPOTPLKGP 3697  
QY 92 ----- 91  
Db 3698 AVEALDSEGEIATFVKOTILKAALITEENQOLSIEHIAKANELSSQLPLGAELQSLIEQ 3757  
QY 92 ----- 91  
Db 3758 DKLTPESTREFLCINGSIHQPLKEPSPNLOLQIVQSOKTFSEKGIILMPPEPETOAVLSD 3817  
QY 92 ----- 91  
Db 3818 TEKIFPSAMSTEQINSLTVEPLKTLAEPEGNYPOSSIEPPMHSYLTSVAEEVLSPEKT 3877  
QY 92 ----- 91  
Db 3878 VSDTNREORVTLOKEAQSALILQSIAEGHVESLOSQPDVNI SQVNYEPLVPSEHSCTEG 3937  
QY 92 ----- 91  
Db 3938 GKILLIESANPLENAGDSAVRIEBSKSLRPLALEEKQVLLKEHSDNVVWPPDQIIIESK 3997  
QY 92 -----QDWL- 95  
Db 3998 REPVAIKKVOEVOGRDILLSKESLLSGIPERQRLNLKIQICRALQAQAAVASEQPLFSEWL 4057  
QY 96 ----- 95  
Db 4058 NIEKVEAVNITQEPHIMCMYLVTSKASVTEEVTHIIEDVDPQMANLKWELRDALCAI 4117  
QY 96 ----- 95  
Db 4118 IYEEIDILTAEGPRIQOQAKTSLQEENDSFGSKQKVEPITEPEVESKYLISPEEVSFNV 4177  
QY 96 ----- 95  
Db 4178 QSRVKYLDATPVTKGASAVVDEKODESLKPSKEEBSSESGETEVATVKIQEABGF 4237  
QY 96 ----- 95  
Db 4238 IKEDGPMIHTPLVDTVSEGDIVHLTTSITNAKEVNWYFENKLVPSDEKPKCLQDQNTYT 4297  
QY 96 -----NGK----- 98  
Db 4298 LVIDKVNTEDHQGEYVCEALNDSKATSAKLTVVKRAAPVKKRIEPLVALGHLAKFT 4357  
QY 99 -----EYKCKVSN----- 106  
Db 4358 CEIOSAPNVRQWFKAGREIYESDKCSIRSKYISSLEILLRTQVDCGEYTCRASNEYS 4417  
QY 107 ----- 106  
Db 4418 VSCATLTVTPGGEKKVKRLPERKPEKEEVVLSVKLRPERPEEPKVEPKKLEKVK 4477  
QY 107 ----- 106  
Db 4478 PAVPEPPPPKVEVEVPTVKRERKIPEPTKVPKPAIPKPAPEPKPPEAEVKTIPK 4537  
QY 107 -----KALPAPIE----- 114

Db 4538 PPVEPEPTPIAAPVTVPVVGKAEAKAPKEAAKPGIKGVPKKTPSIEAERRKLPG 4597  
QY 115 ----- 114  
Db 4598 SGGEKPPDEAFFTYQLKAVPLKFVKEIKDIIILTESEFVGSSAIFECVLSPSTAITTMMKD 4657  
QY 115 ----- 114  
Db 4658 GSNIRESPKHFRIADGDKDKLHIIIDVQLSDAGEYTCVLRGNKEKSTAKLVVEELPVRF 4717  
QY 115 -KTIS----- 118  
Db 4718 VKTLEEVTVVKGQPLYLSCNELKRDVWRKDGKIVVEKPGRIVPGVIGLMRALTINDA 4777  
QY 119 ----- 118  
Db 4778 DDTDAGTYTVTVENANNLECSVCVKVVEVIRDMLVKPIRDQHVKPKGTAFACDIADKTP 4837  
QY 119 ----- 118  
Db 4838 NIKWFKGYDEIPAPBNPKTEILRDGNHLYLKIKNAMPEDIAEYAVEIEGKRYPAKLTLGE 4897  
QY 119 ----- 118  
Db 4898 REVELLKPIEDVTIYKESASFDAEISEADIPGOWKLKELLRPSPTCEIKAEGGKRFLT 4957  
QY 119 ----- 118  
Db 4958 LRKVKLDQAGEVLYQALNAITTAITLVKEIELDFAVPLKDVTPPERQARFECVLTREAN 5017  
QY 119 ----- 118  
Db 5018 VIKSGPDIILKSSDKPDIADGKKHILVINDSQFDDGCVYTAEVGKKT SARLFTVGTGIRL 5077  
QY 119 ----- 118  
Db 5078 KFMSPLEDQTVKEGETATFVCELSHERKMHVWFKNDAKLHTSRTVLISSEKTHKLEMKE 5137  
QY 119 ----- 118  
Db 5138 VTLDISQIKAQVKELSSSTAQLKVLEADPYFTVKLHKTAVEKDEITLKCEVSKDVPVKW 5197  
QY 119 ----- 118  
Db 5198 FKDGEEIVPSPKYSIKADGLRRLIKKADLKDKEVYVCDGCTDKTKANVTVEARLIKVE 5257  
QY 119 -----KAGQP----- 124  
Db 5258 KPLYGVEFVGETAHFEIELSEPDVHGQWKLKGPLTASPDCEIIEDGKKHILILHNCQL 5317  
QY 125 -----REPQ----- 128  
Db 5318 GMTGEVSFOAANAKSAANLKVKELPLIFITPLSDVKVFEXDEAKFECEVSREPRTFRWLK 5377  
QY 129 ----- 128  
Db 5378 GTQEITGDDRFEILKDGTKHSMVIKSAAFEDKAYMFEADKHTSGKLIIEGIRLKLFTP 5437  
QY 129 ----- 128  
Db 5438 LKDVTAKEKESAVFTVELSHDNIRVKWFKNDQRLHTTRSVSMQDEGKTHSITFKDLSIDD 5497  
QY 129 ----- 128  
Db 5498 TSQIRVEAMGMSSEAKLTVLEGDPYFTGKLODYTGVEKDEVILQCEISKADAPVKWPKDG 5557  
QY 129 ----- 128  
Db 5558 KEIKPSKNAVIKADGKKRMLILKALKSDIGQYTCDCGTDKTSCKLDIEDREIKLVRPLH 5617  
QY 129 ----- 128

Db 5618 SVEVETETARFETEISEDDIHANWKLGEALLQTPDCEIKERKTHLSVLHNCRLDQTG 5677  
QY 129 ----- 128  
Db 5678 GVDFOANVKSSAHLRVKPRVIGLLRPLKDVTVTAGETATFDCELSYEDIPVEWYLGKK 5737  
QY 129 ----- 128  
Db 5738 LEPSDKVPRSEKGVHTLRLRDVKLEDAGEVQLTAKDFKTHANLFVKEPPVEFTKPLEDQ 5797  
QY 129 ----- 128  
Db 5798 TVEGATAVLECEVSRENKRVFKNGTEILSKKYEIVADGRVKLVIHDCTPEDIKTY 5857  
QY 129 ----- 128  
Db 5858 TCDAKDFKSCNLNVPPHVEFLRPLTDLOVREKEMARFECELSRENKRVKFKDGAETK 5917  
QY 129 ----- 128  
Db 5918 KGKKYDIISKGAVRIILVINKCLLDDAEYSCEVRTARTSGMLTVLBEAEVFTKNLANIEV 5977  
QY 129 ----- 128  
Db 5978 SETDTIKLVEVSKPGAIEVIWYKGDEEIIETGRYEILTEGRKRIILVIONAHLEDAGNYC 6037  
QY 129 ----- 128  
Db 6038 RLPSSRTDGKVKVHELAAEFISKQONLEILEGEKAEFVCSISKESFPVQWKRDDKTLESG 6097  
QY 129 ----- 128  
Db 6098 DKYDVIADGKKRVLVVKDATLODMGTVMVVGAAARAHAHLTVIEKLRIVVPLKDTFVKQ 6157  
QY 129 ----- -VYTL----- 132  
Db 6158 QEYVFNCEVNTGAKAKWPNBEAIPDSSKYIILOKDLVYTLRIIRDAHLDDQANVSLT 6217  
QY 133 ----- 132  
Db 6218 NHRGENVKSANLIVEEDLRIVEPLKDIETWEKKSVTFWCKVNRNLNVLTKWTKNGEVP 6277  
QY 133 ----- 132  
Db 6278 FDNRSYRVDKYKHMLTIKDCGFPDGEYIVTAGQDKSVAELLIIEAPTEFVEHLEDQTV 6337  
QY 133 ----- 132  
Db 6338 TEFDDAVFCQLSREKANKWYRNGREIKEGKKYKFEKDGSIHRLIHKDRLDDECEYAC 6397  
QY 133 ----- 132  
Db 6398 GVEDRKSARLFVEETPVEIIRPPQDILEAPGADVVFLELNKDKVEQWLNRNMVVOG 6457  
QY 133 ----- 132  
Db 6458 DKHOMSEKIHRLQICDIKPRDGEYRIAKDKARAKLELAAPKIKTADQDLVDVG 6517  
QY 133 ----- 132  
Db 6518 KPLTWVPYDAYPKAAEWFKENEPLSTKTIDTTAOTSFRILEAKGDKGRYKIVLQNK 6577  
QY 133 ----- 132  
Db 6578 HGKAEFINLVIDVPGVPVRNLEVTFDGEVSLAWEEPLTDGGSKIIGYVVERDIRK 6637  
QY 133 ----- 132  
Db 6638 TWLATDRAESCEFTVTGLQGGVEYLFVRARNRVGTGEPVETDNPEARSKYDVPGP 6697  
QY 133 ----- 132  
Db 6698 LNVITDVRNFGVSLTWPEPYDGGAEITNVVIELRDKTSIRWDTAMTVRAEDLSATVTD 6757

QY 133 ----- 132  
Db 6758 VVEGOEYSPRVAQNRIQVCKPSAATPFVKVADPIERPSPVNLITSSDOTQSSVQLKWP 6817  
QY 133 ----- 132  
Db 6818 PLKGGSPILGYIIERCEBGNWIRCNMKNLVPULTYKVTGLEKGNKYLYRVAENKAGV 6877  
QY 133 ----- 132  
Db 6878 SDPSEILPLTADDAFVEPTMDLSAFKDGLEVIVPNPITILVPSTGYPRPTATWCFGDKV 6937  
QY 133 ----- 132  
Db 6938 LETGDRVKMKTLSAYAEVLISPSERSDKGIYTLKLENRVKTTISGEIDVNVIAIAPSAPKEL 6997  
QY 133 ----- 132  
Db 6998 KFGDITKDSVHLTWBPPDDGGSPLTGYVVEKREVSRTWTWKVMDFTDLEFTVPLVQ 7057  
QY 133 ----- 132  
Db 7058 KEYLFKVCARNKCGPEPAYVDEPNMSTPATVPDPENVKWRDRTANSIFLTWDPKND 7117  
QY 133 ----- 132  
Db 7118 GGSRIKGYIVERCPRGSKWVACGEPVATKMEVTGLEBEGKYAYRVKALNRQGASKPSR 7177  
QY 133 ----- 132  
Db 7178 PTEEIQAVDTQEAPEIFLDVLLAGLVKAGTKIELPATVTGKPEKTIWTKADMILKQD 7237  
QY 133 ----- 132  
Db 7238 KRITIENVPKSTVTIVDSKRSDTGYIIIEAVNVCGRATAVVEVNVLDKPGPPAAFDITD 7297  
QY 133 ----- 132  
Db 7298 VTNESCLLTWNPDRDDGSKITNVYVVERRATSEVWHKLSSTVKTNFKATKLIPNKEVI 7357  
QY 133 ----- 132  
Db 7358 FRVAENMVGCEPVOASPIAKYQFDPGPPTRLEPSDITKDAVTLTWCEPDDGGSP 7417  
QY 133 ----- 132  
Db 7418 TGYWYERLDPDTKWVRNKNMPVKDTTVRVKGLTNKKYRFRVLAENLAGPKSPKSTEP 7477  
QY 133 ----- 132  
Db 7478 ILIKDPIDPPPGKPTVKDVCKTSVRLNWTXPEHGGAKIESYVIEMLKTGTDEWVRVA 7537  
QY 133 ----- 132  
Db 7538 EGVPTTQHLPLGMEGQEVYFRVAVNKAAGESEPSPVLCREKLYPPSPRWLEVIN 7597  
QY 133 ----- 132  
Db 7598 ITKNTADLKWTVPKDGGSPTNYIVEKRDVRRKGQTVDTTIVKDTCTVTPLTEGSLYV 7657  
QY 133 ----- 132  
Db 7658 FRVAENAIGSDYTEIEDSVLAKDTFTTPGPPYALAVVDYTKRHVDLKWEPKNDGGPR 7717  
QY 133 ----- 132  
Db 7718 IQRYVIEKKERLGRWVAGKTAGPCNFRVTDVIEGTEVQFQVRAENAGVHPSEPTE 7777  
QY 133 ----- 132  
Db 7778 ILSIEDPTSPSPPLDLHLVTDAGRKHIAIAWKPPBKNGSGSPIIGYHVEMCPVGTCKMRV 7837



QY 133 ----- 132  
Db 7838 NSRPIKDLKFVVEGWDPKEYLVRRAVNAIGVSEISENVVAKDPCKPTIDLETH 7897  
QY 133 ----- 132  
Db 7898 DIIVIEGEBKLSIPVFRAPVPTVSMHKDGKEVKASDRLTMKNDHISAHLEVPKSVRADA 7957  
QY 133 ----- 132  
Db 7958 GIYTITLENKLSATASINVKVIGLPGCKDIKASDITKSSCKLTWEPEDGGTPILHY 8017  
QY 133 ----- 132  
Db 8018 VLEREAGRRTYIPVMGENKLSWTVKDLIPNGEYFFRVKAVNKVGGGEYIELKNPVIAQ 8077  
QY 133 ----- 132  
Db 8078 DPKQPPDPVDEVNHTAEAMTITWKPLDYDGGSKIMGYIIEKTAKGEERWKRCNEHLV 8137  
QY 133 ----- 132  
Db 8138 PILTYTAKLEGEKEYQFVRRAENAGISEPSRATPPTKAVDPIDAPKVILLRTSLEVKRG 8197  
QY 133 ----- 132  
Db 8198 DEIALDASISGSPYPTITWIKDENVIVPEEIKKRAAPLVRRRKGEVQEEEPVLPLOTQL 8257  
QY 133 ----- 132  
Db 8258 SIDSKGESQLVRDSLRDPHGLMIKVENDHGIAKAPCTVSVLDTFGPPINFVEDIR 8317  
QY 133 ----- 132  
Db 8318 KTSVLCKWEPPLDGGSEIINTYLEKDKTKPDSEMIWVTSTLRHCKYSVTKLIEGKEYL 8377  
QY 133 ----- 132  
Db 8378 FRVRAENRFGPGPCVSKPLVAKDPGPPADPKPIVEDVTSNMLVKWNEPKONGSPIL 8437  
QY 133 ----- 132  
Db 8438 GYWLEKREVNTHSRVKNKSLNALKANVDGLLEGLTYVFRCAENAAAGPGKFSPPDPK 8497  
QY 133 ----- 132  
Db 8498 TAHDPISPGPIPRVTDTSSTTIELEWEPFANGGEIYGVYVDKQLVGTNEWSRCTEK 8557  
QY 133 ----- 132  
Db 8558 MIKVRQYTVKEIREGADYKLRVSAVNAAGEGPGGETQPVTVAEPOEPPAVELDVSVKGI 8617  
QY 133 ----- 132  
Db 8618 QIMAGKTLRIPAVTGRPVPTKVWTKKEGELDKRVIDNVGTSKELI IKDALRKDHGRY 8677  
QY 133 ----- 132  
Db 8678 VITATNCSGKFAAARVEVDVPGVPYLDLKPVVTNRKMCLLNWSDEDDGGSEITGFIIE 8737  
QY 133 ----- 132  
Db 8738 RKDAKHTWROIETERSKCDITGLLEGQYKFRVIANKFGCGPPVEIGPILAVDPLGP 8797  
QY 133 ----- 132  
Db 8798 PTSPELRTVTKSTITLWDKEPRNSGGSPIQGYIIEKRRHKDPDFERNKRLCPTTSF 8857  
QY 133 ----- 132  
Db 8858 LVENLDEHOMYEFVRKAVNEIGESEPSLPINVIQDDEVPPTIKURLSVRGDTIKVKAGE 8917  
QY 133 ----- 132

Db 8918 PVHIPADVTLGPMPIKIEWSKNETVIEKPTDALQITKEEVSSEAKTELSIPKAVREDKGT 8977  
QY 133 ----- 132  
Db 8978 YTVTASNLGSGFRNVHVEYDRPSPRNLAUTDIKASCYLTWAPLDNGGSEITHYVI 9037  
QY 133 ----- 132  
Db 9038 DKRDASRKAWEWEVTNTAVEKRYGIWKLI PNGQYEFVRVAVNKYGISDECKSDKVVIQD 9097  
QY 133 ----- 132  
Db 9098 PYRLPGPCPKVLARTKSMVLSWTPPLDNGGSPITGYWLEKREBGSYWSRVSRAPIT 9157  
QY 133 ----- 132  
Db 9158 KVGLKGVEFVPRLLLEGVYQFRAMAINAAGIGPSEPSDPEVAGDPIFPFGPPSPCEVK 9217  
QY 133 ----- 137  
Db 9218 DKTSSISLGWKPAPKDGSPKIGYIVEMQEEGTTDWKRVNBDKLIITTCCECVPNLXEL 9277  
QY 138 ----- 137  
Db 9278 RKYFRVRAVNEAGESEPSDITTEIPATDIOEEPEVFI DIGAODCLVCKAGSQIRIPAVI 9337  
QY 138 ----- 137  
Db 9338 KGRTPKSWEPDGGKAKKAMKGVHDIPEDAOETAENSSVLIIECKRSHGTGKYSITAK 9397  
QY 138 ----- 137  
Db 9398 NKAGKTANCVRKVMVDPGPPKDLKVS DITRGSCRLSWKMPDDGGDRIGYVIEKRTID 9457  
QY 138 ----- 137  
Db 9458 GKAWTKVNPDCGTTTFVVPDLLSEQYFFRVAENRFGIGPPVETIQTARTDPIYPPDP 9517  
QY 138 ----- 137  
Db 9518 PIKLKIGLITKNTVHLSWPKPKNDGSGPVTHYIIVECLAWDPTGTKEAWRQCNKRDEEL 9577  
QY 138 ----- 137  
Db 9578 QFTVEDLVEGGEYEFVRKAVNAAGVSKPSATVPCDCQRPDPSPSIDLKEFMEVEBGTNV 9637  
QY 138 ----- 137  
Db 9638 NIVAKIKGVPPFTLTFWKAPKKPNKEPVLVDYTHVKNLVDDTCTLVIPQSRSSDTGLY 9697  
QY 138 ----- 137  
Db 9698 TITAVNNLGTASKEMRLNVLRGPPVPGDIKFSVSADQMTLSWPPPKDDGSKITNYVI 9757  
QY 138 ----- 137  
Db 9758 EKREANRKTWVHVSSEPKECTYTI PKLLEGEHYVFRIMAQNKYGIGEPDSEPETARNLF 9817  
QY 138 ----- 137  
Db 9818 SVPGAPDKPTVSSVTRNSMTVNWEEPEYDGGSPVTGYWLEMKDOTTSKRWKRVNRDPIKAM 9877  
QY 138 ----- 137  
Db 9878 TLGVSKYKTGLIEGSDYQFRVYAINAAGVGPASLPSPATARDPIAPGPPPKVTDWTK 9937  
QY 138 ----- 137  
Db 9938 SSADLEWSPPLKDGSKVTGYIYEYKEEGKEWEKDKVEVRCTKLVVTGLKEGAFYKFR 9997  
QY 138 ----- 137

```

Db 9998 VSAVNIAGIGEVEYTDVEMKDLVSPDLQLDASVRDRIVVHAGVIRIIAYSGKPPP 10057
QY 138 ----- 137
Db 10058 TVTWMNERTLPOEATIIETAISSMVKINCORSHQGVYSLAKNEAGERKKTIIIVDLD 10117
QY 138 ----- 137
Db 10118 VPGPVGTPLAHLNLTNESCCLTWFSPEDDGSPITNYVIERESDRRAWTPTVTVTRQN 10177
QY 138 ----- ELTKNOVS 145
Db 10178 ATVGLIOGKAVFFRIAAENSGMGPFVETSEALVIREPIVPERPEDELVKEVTKNTVT 10237
QY 146 LT----- 147
Db 10238 LTWNPPKYDGGSEIINYLESRLIGTERFHKVTNDNLLSRKYTVKLGEGDTYEYRSVA 10297
QY 148 ----- 147
Db 10298 NIVGQKPSFCTKPTCKDELAPTLHLDFRDKLTIRVGEAFALTGRYSRKPKVSWFK 10357
QY 148 -----CLV-----KGF----- 153
Db 10358 DEADVEDDRTHIKTTPATLALEKIKAKRSDSGKYCVVENSTGSRKGFCQVNVVDRGP 10417
QY 154 ----- 153
Db 10418 PVGPSEFVTKDYMVISWKPPLDGGSKITNYIIERKEVGKDVMPVTSASAKTTCKVS 10477
QY 154 ----- 153
Db 10478 KLEBGDYIPRIHAENLYGSDPLVSDSMKAKDRFRVPDAPDQPIVTEVTKDSALVTWNK 10537
QY 154 ----- 153
Db 10538 PHDGGKPTINYILEKRETMKRWARVTKDPIHTYTKFRVPDLLBGCOYEFVSAENEIGI 10597
QY 154 ----- 153
Db 10598 GDSPPSKPVFAKDPIAKSPVNPENAIIDTCSNVDLTWQPRHDGSKILGYIYEQV 10657
QY 154 ----- 153
Db 10658 GDEWRANHTPESCPETKYVTCLRDQYTKFRVLAVNAAGESDPAHVPEPVLVDRLE 10717
QY 154 ----- 153
Db 10718 PPELILDANMAREQHIKVGDTLRLSAIKGVFPKVTWKEDRDAPTARIDVTPVGSKL 10777
QY 154 ----- 153
Db 10778 EIRNAHEDGGIYSLTVENPAGSKTVSVKVLVDKPGPPRDLVSEIRKDSCLTWKEPL 10837
QY 154 ----- 153
Db 10838 DGGSVITNYVVERRDVASAOWSPLSATSKKSHFAKHLNEGQYLFRVAENQYGRGPF 10897
QY 154 ----- 153
Db 10898 VETPKIKALDPLUHPGPKDLHHVDVDTESVLVWNKPDROGGSPTIGYLVEYQEGTQ 10957
QY 154 ----- 153
Db 10958 DWIKFXTVNLGVVTLQOGKTYRFRVKAENIVGLGLDPTTPIECQKLVPPSPVELDV 11017
QY 154 ----- 153
Db 11018 KLIEGLVWAGTVRFPAILIRGVPVPTAKWTTDGSSEIKTDEHYTVETDNFSSVLTIKNCL 11077
QY 154 ----- 153
Db 11078 RDTGEYQITVSNAGSKTVAVHLTVLDVPGPPTGINILDVTPHEMTISWQPPKDDGS 11137

```

```

QY 154 ----- 153
Db 11138 PVINYIVEKQDTRKDTWGVVSSGSSKTKLIPHLOKGEYFVRVAENKIGVPPDSTP 11197
QY 154 ----- 153
Db 11198 TVAKHKFSPSPGPKPVVTDITENAAVSWTLPKSDGSGPITGYMYERREVTKGWVRNK 11257
QY 154 ----- 153
Db 11258 TPIADLKFRVTLGVEGNTYEFVFAENLAGLSKSPSPDIKACRPIKPPGPPINPKLKD 11317
QY 154 ----- 153
Db 11318 KSRETADLVWTKPLSDGSGPILGYVVECOKPGTAQWNRINKDELIRQCAFVRPGLIEGNE 11377
QY 154 ----- 153
Db 11378 YRFRIKAANIVGEGEPRELAESVIAKDIILHPPEVELDVTCRDVITVRVGQTIRILARVK 11437
QY 154 ----- 153
Db 11438 RPEPDITWTKREGKVLVREKRVDLIQDLPRVELOIKEAVRADHGKYIIISAKNSGHAQSA 11497
QY 154 ----- 153
Db 11498 IVNVLDPRGPCQNLKVNTVKENCTISWENPLDNGSGSEITNFIVEYRKPNOKGHSIVASD 11557
QY 154 ----- 153
Db 11558 VTKRIKANLANNEYFVCAENKVGVPETIETKPTILAINPIDRPEPENLHADGK 11617
QY 154 ----- 153
Db 11618 TFVYLKWRPDPYDGGSPNLSTYHVHERRLUKSGDDWERVHKGSIKETHYMWDRVCVENQIYEFR 11677
QY 154 ----- 153
Db 11678 VQTKNEGGEDWVKTEBWWVVKEDLQKPVLDKLGLSVLTVKAGDTIRLEAGVRGKPFPEVA 11737
QY 154 ----- 153
Db 11738 WTKDKDATDLTRSPRVKIDTRADSKFSLTAKAKRSDGGKYVVVATATAGSFVATVNVL 11797
QY 154 ----- 153
Db 11798 DKPGPVRLKIVDVSSDRCTVCHDPPEDDGGCEIQNYILEKCETKRMVWSTYSATVLTGP 11857
QY 154 ----- 153
Db 11858 TTVTRLIEGNEYIFRVAENKIGTPPTESKPVIAKTKYDKPGRPDPPEVTKVSKBEMTV 11917
QY 154 ----- 153
Db 11918 VWNPPYDGGKSTIGTYFLEKKEKHSTRVWPVNKSAIPERRMKYONLLPDHEYQFVKAEN 11977
QY 154 ----- 153
Db 11978 EIGIGESLPSRPVAKDPIEPPGPTNFRVVDTTKHSITLGHGKPYDGGAPIIGYVE 12037
QY 154 ----- 153
Db 12038 MRPKIADASPDEGWKRCNAAQLVRKEFTVTSLEDNEQEYFVCAQNOVGIGRPAELKEA 12097
QY 154 ----- 153
Db 12098 IKPKEILEPPEIDLDASMRKLVIVRAGCPIRLFAIVRGRPAKVTWRKVGIDNVVRKGQV 12157
QY 154 ----- 153
Db 12158 DLVDVTMAFLVIPNSTRDDSGKSLTLVNPAGEKAFVNVVRVLDTPGVSCLKVSDVTKTS 12217

```

QY 154 ----- 153  
Db 12218 CHVSWAPPENDGGSQVTHYIVKREADRTWSTVTPVKKTSFHTNLNPGNEYFRVTA 12277  
QY 154 ----- 153  
Db 12278 VNEYGCVPTDVKPVLASPLSEDPDRKLEVTMTKNSATLAWPLPLRGGAKIDGYI 12337  
QY 154 ----- 153  
Db 12338 TSYREEQPADRWTEYSVVKDLSLVITGLKEGKKYKFRVAARNAGVSLPREAEGVYEAK 12397  
QY 154 ----- 153  
Db 12398 EQLPPRIILMEQITIKAGKRLIEAHVYKPHPTCKWKKEDEVVTSSHLAVHKADSS 12457  
QY 154 ----- 153  
Db 12458 ILIKDVRKDSGYYSUTAENSSGTDQKIKVVVMDAPGPPQPPFDISDIDADACLSWH 12517  
QY 154 ----- 153  
Db 12518 IPLEDGGSNITNYIVEKCDVSRGDWVTALASVTKTSRVCGLIPGOEYIFRVAENRFGI 12577  
QY 154 ----- 153  
Db 12578 SEPLTSPKMAQPPFGVPSEPKNARVTKNKDCIFVANDRPDSGGSPILIGYLIERKERN 12637  
QY 154 ----- 153  
Db 12638 SLLWVKANDTLVRSTEYPCAGLEVEYSFRIYALNKAGSPSPKTEYVTARMPVDPG 12697  
QY 154 ----- 153  
Db 12698 KPEVIDVTKSTVSLIWARPKHDGSKLIIGYFVEACKLPGDKWRCNTAPHQIQEETAT 12757  
QY 154 ----- 153  
Db 12758 GLEKAQYQRAIARTAVNISPPSPDPVTILAENVPRIIDLVSAMKSLITVKAGTNVC 12817  
QY 154 ----- 153  
Db 12818 LDATVFGKPMPTVSWKDGTLKPAEGIKMAMORNLTLEFSVNRKDSGDTITAENSS 12877  
QY 154 ----- 153  
Db 12878 GSKSATIKLVLDKPGPPASVKINKMYSDRAMLSWEPPLEDGGESEITNYIVDKRETSRPN 12937  
QY 154 ----- 153  
Db 12938 WAQVSATVPTITSCSVKLEIEGHEYQFRIKAENKYGVGDPVFTTEPAIAKNFYDPPGRCDPP 12997  
QY 154 ----- 153  
Db 12998 VISNITKDHMTVSWKPPADGGSPITGYLLEKRETOAVNWTKVRKPIIBERTLKATGLQE 13057  
QY 154 ----- 153  
Db 13058 GTEYFRVTAINKAGPCKPSDASKAAYARDPQPPAPPAPPKYVDTRSSVLSWGPAY 13117  
QY 154 ----- 153  
Db 13118 DGGSPILIGYLVKRADSDNWRNLPQNLQKTRFVTEGLMEDTQYQFRVAVNKGYS 13177  
QY 154 ----- YPSDIIV  
|| || : 160  
Db 13178 PSDVPDKHYPKDILIPPEGEDADRLKTLILRAGVTMRLYVPVKGRPPPKITWSKPNVL 13237  
QY 161 ----- 160  
Db 13238 RDRIGLDIKSTDDFTFLRCENVNKYDAGKYILILENSCGKKEYTIVVKVLDTPGPPNVNT 13297  
QY 161 ----- 160

Db 13298 VKEISKDSAYVTWEPPIIDGGSPIIINYVQKRDAERKSWSTVTECSKTSFRVANLEEGK 13357  
QY 161 ----- 160  
Db 13358 SYFERVFAENEYIGIDPGETRDVAKASQTPGVVDLKVRSYSKSSCSIGWKKPHSDGGS 13417  
QY 161 ----- 160  
Db 13418 IIGVVDLFTENKQWRVMSLSLQYSAKDLTEGKEYTFRVSAENEGEPTSEITVVAR 13477  
QY 161 ----- 160  
Db 13478 DDVAPDLDLKGLPDLCLYLAKENSFRLKIPKIGKPAPSVWKKGEDPLATDTRVSSESS 13537  
QY 161 ----- 160  
Db 13538 AVNTTLIVDCQSDAGKYITILKNVAGTKEGTISIKVVGKPGIPTGPIKFEVTAEMT 13597  
QY 161 ----- 160  
Db 13598 LKWAPPKDDGSEITNYILEKRDSVNKNKWTCASAVQKTTFRVTRLHEGMEYTPRVSAN 13657  
QY 161 ----- 160  
Db 13658 KYGVGGLKSEPIVARHPFDVDPAPPPNIVDVHRDSVSLTWTDPKKTGGSPITGYHLEF 13717  
QY 161 ----- 160  
Db 13718 KERNSLLWKANKTPIRMROFKVTGLTEGLEYEFRVMAINLAGVKFSLPSEPVALDPI 13777  
QY 161 ----- 160  
Db 13778 DPGKPEVINITRNSVTLIWTEPKYDGHKLTGYIVEKRDLPKSKWMKANHVNPECAFT 13837  
QY 161 ----- 160  
Db 13838 VTDLVEGGYEFIRAKNTAGAISAPSESTETIICKDEYAPTIVLDPITKDGLTIRAGD 13897  
QY 161 ----- 160  
Db 13898 TIVLNAISILGKPLPKSSWSKAGKDIRPSDITQITSTPTSSMLTIKYATRKDAGEYITA 13957  
QY 161 ----- 160  
Db 13958 TNPFGTKVHVKTIVLDVPGPPGVEISNVSAAKATLTWTTPLEDGGSPIKSYILEKRET 14017  
QY 161 ----- 160  
Db 14018 SRLLTWTVSEDIQSRHVATKLIQONEYIFRVSANHYGKGEVQSEPVKMDRFGPPGP 14077  
QY 161 ----- 160  
Db 14078 PEKPEVSNVTNTATVSWKRPVDDGGSEITGYHVVERREKSKLRVRAIKTPVSDLRCKVT 14137  
QY 161 ----- 160  
Db 14138 GLOGESTYFRVSAENRAGIGPSEASDSVLMKDAAYPPGPPSPNPHVTDITTKSASLAWG 14197  
QY 161 ----- 160  
Db 14198 KPHYDGGLEITGYVVEHQVKGDEAWIKDQTTGTLRITQFVVVDLQTKENYFRISAINDA 14257  
QY 161 ----- 160  
Db 14258 GVGEPAVIPDEIVEIREMAPDFELDAELRRLTVVRAGLSIRIFVPIKGRPAPEVTWKDN 14317  
QY 161 ----- 160  
Db 14318 INLKNRANIENTESFTLLIIPECNRYDTGKFVMTIENPAGKSGFVNVRVLDTPGPNL 14377  
QY 161 ----- 160

Db 14378 RPTDITKDSVTLHWDLPIDGGSRIITNYIVEKREATRKSYSSTATTKCHCKTYKVTLGSEB 14437  
QY 161 ----- 160  
Db 14438 CEYFFRMAENEXGIGETETPEVKASEAPSPDLSNIMDITKSTVSLAWPKPKHDGGS 14497  
QY 161 ----- 160  
Db 14498 KITGYVIEAQRKSGDQWHTITTVKGLCVVRNLTEGEEYTFQVMVNSAGRSAPRESRV 14557  
QY 161 ----- 160  
Db 14558 IVKEQTMPELDLRGIYQKLVIAKAGDNKIVEIPVLRPKPTVTKKGDIILKQTORVNF 14617  
QY 161 ----- 160  
Db 14618 ETTATSTILNINECVSDSGPYPLTARNIVGEVDVITIQVHDIPGPTGPIKFDVSSD 14677  
QY 161 ----- 160  
Db 14678 FVTFSDPPDGGVPISNVVEMRQDSTTWVELATTVIRTTYKATRLTTGLEYOFRVK 14737  
QY 161 ----- 160  
Db 14738 AQNRVGVGPIITSACIVANYPFKVPGPGTPQVATVKDSMTISMHEPLSDGGSPIILGYH 14797  
QY 161 ----- 160  
Db 14798 VERKERGILWQTVSKALVPGNIFKSSGLTDGIAEFVRVIAENMACKSKPSKPSBMLAL 14857  
QY 161 ----- 160  
Db 14858 DPIDPPKVPPLNITRHTVTLKAKPEYTGFKITSYIVEKRDLPNGRWLKNFNSILEN 14917  
QY 161 ----- 160  
Db 14918 EFTVSLGTEDAAEFVRVIAKNAAGAISSPSPSDAICTRDDVEAPKIKVDVRFKDTVILK 14977  
QY 161 ----- 167  
Db 14978 AGEAFLEADVSRPPPTMEWSKDGKELGTAKLEIKIADFSTNLVKNKSTRRDSCAYTL 15037  
QY 168 ----- 167  
Db 15038 TATNPGGFAKHIFNVKVLDRPGPEGLAVTEVSEKCVLSWFPPLDGGAKIDHYIVQK 15097  
QY 168 ----- 167  
Db 15098 RETSLRLANTVASEVOVTKLVTKLLKGNEYIFRVMVNVKYGVEPLESEPVLAVPNYPG 15157  
QY 168 ----- 167  
Db 15158 PDPPKNPEVTTITKDSMVVWGHDPDGSSEIINIYVERDRKAGORWIKCNKKTLDLRY 15217  
QY 168 ----- 167  
Db 15218 KVSGLTEGHEYFRIMAENAGISAPSPSPFYKACDITVFKPGPPGNPRVLDTSRSSISI 15277  
QY 168 ----- 167  
Db 15278 ANKPIYDGGSEITGYMVEIALPEDEWOIVTPPAGLKATSYTITGLTENQEKIRIYAM 15337  
QY 168 ----- 167  
Db 15338 NSEGLGEPALVPCTPKAEDRMLPPELELDADLRKVVTIRACCTLRFLFPIKGRPAPEVKW 15397  
QY 168 ----- 167  
Db 15398 ARDHGESLDKASTESTSSYTLIVGNVRNFRDGSKVILAVENSSGSAFVNVRLDTPG 15457  
QY 168 ----- 167  
Db 15458 PODLKVETKTSVTLTWDPPLDGGSKIKNYIVEKRESTRKAYSTVATNCHKTSWKVDQ 15517

QY 168 ----- 167  
Db 15518 LQEGCSYFVRVLAENEYIGLPAETAESVKASERPLPPCKITILMDVTRNSVLSWEKPEH 15577  
QY 168 ----- 167  
Db 15578 DGGSRILGYIVEMQTKGSKWATCATVKVTEATITGLIQGEEYSFRVSAQNEKIGSDPRQ 15637  
QY 168 ----- 167  
Db 15638 LSVPIAKDLVIPPAPKLLFNTFTVLAGEDLKVDVPFIFGRPTPAVTHKDNVPLKQTTVR 15697  
QY 168 ----- 167  
Db 15698 NAESTENNSLLTIKDACREDVGHYVVKLTNSAGEAETLNVLDPKPGPTGPKVMDEVY 15757  
QY 168 ----- 167  
Db 15758 ADSITLSWGPYPKYDGGSSINNYIVEKRDSTTTTQWIVSATVARTTIKACRLKTGCEYQPR 15817  
QY 168 ----- 167  
Db 15818 IAAENRYGKSTYLNSEPTVAQYFVKVPGPGTPVVTLSRDSRDSMEVQWNEPISDGGSRVIG 15877  
QY 168 ----- 167  
Db 15878 YHLERKERNISILWVLNKTPIPQTKFTGTLEEGVEYEFVRVSAENIVGIGKPSKYSECV 15937  
QY 168 ----- 167  
Db 15938 ARDPCDPCRPEAIIVTRNSVTLOWKKPTYDGGSKITGYIVEKKELPEGRMWSKASTNII 15997  
QY 168 ----- 167  
Db 15998 DPHFEVTLGVEDHRYEFVRVIAARNAAGVSEPSSESTGATARDEVDPPIRISMDPKYKDTIV 16057  
QY 168 ----- 167  
Db 16058 VHAGESFKVDADYICKPIPTIOWIKGQDQELSNTARLEIKSTDFATSLSVKDAVRVDSGNY 16117  
QY 168 ----- 167  
Db 16118 ILKAKNVAGERSVTYVNVKVLDRPGPEGVVIGSVTAEKCTLAWKPPLODGGSDIINYIV 16177  
QY 168 ----- 167  
Db 16178 ERRETSRLVWTVDANVQTLCKVTKLLEGNEYTFRIMAVNVKYGVEPLESEPVVAKNPF 16237  
QY 168 ----- 167  
Db 16238 VVPDAPKAPEVTTVTKDSMIVVWERPASDGGSEILGYVLEKKRDEGIRWTRCHKRLIGEL 16297  
QY 168 ----- 167  
Db 16298 RLRVTGLIENHDEYFRVSAENAGLSEPPSPAYQKACDPIYKPGPPNPNKVIDITRSSV 16357  
QY 168 ----- 167  
Db 16358 FLWSKPIYDGGCEITQGYIVEKCDVSGEWTMCTPPTGINKTNIEVEKLEKHEYNFRIC 16417  
QY 168 ----- 167  
Db 16418 AINKAGVEHADVPGLIIEEKELEAPDIDLLELRKIIINIRAGGSLRFLVPIKGRPTBV 16477  
QY 168 ----- 167  
Db 16478 KWGVDEIRDAAIIDVTSSFTSLVDNWNRYDSGKTLTLENSSGTSKSAFVTVRLDTP 16537  
QY 168 ----- 167  
Db 16538 SPPVNLKVTEITKDSVSTWBPPLDGGSKIKNYIVEKREATRKSAAVAVTNCNHSWKI 16597

QY 168 ----- 167  
Db 16598 DQLEGCYYFRYTAENEYIGLPAQTADPIKVAEPQPPGKITVDDVTRNSVLSWTKP 16657  
QY 168 ----- 167  
Db 16658 EHDGSKIIOYIVEMOAKHSEKSECARVKSQAQVITNLTOGEYLFRRVAVNEKGRSDP 16717  
QY 168 ----- 167  
Db 16718 RSLAVPIAKDLVIEDVRKPAFSSYSVOVGQDLKIEVPISGRPKPTITWTKDGLPLKQTT 16777  
QY 168 ----- 167  
Db 16778 RINWTDLDLTLISIKETHKDDGGQYITVANVVGQTASIEIVTLDKDPKPGVKFDD 16837  
QY 168 ----- 167  
Db 16838 VSAESITLSNPPLYTGGCOITNYIVQKRDTTTTVWDVVSATVARTTLKVKLKLTGTEYQ 16897  
QY 168 ----- 167  
Db 16898 FRIFAENRYCOSFALESDDPIVAQPYKPEPGPGTPTAFALSKDSMVIOHHEPVNNGSPV 16957  
QY 168 ----- 167  
Db 16958 IGYHLERKERNLSLWTKVKNKTIIDTQFAQNLEEGIEYEFVRYAENIVGVGKASKNSEC 17017  
QY 168 ----- 167  
Db 17018 YVARDPCDPPGTPPEIMVNRNEITLQWTKPVYDGGSMITGYIVEKRDLPDGRWMKASF 17077  
QY 168 ----- 167  
Db 17078 VIETQTVSGLTEDQRYEFVRIAKNAAGALSKPDSSTGPTAKDEVELPRISMDPKRDT 17137  
QY 168 ----- 167  
Db 17138 IVNAGETFRLEADVHGKPLPTIEMLRGDKEIEBSARCEIKNTDFKALLIVKDAIRIDGG 17197  
QY 168 ----- 167  
Db 17198 QYILRASNAVAGSFPVNVKVLDRPDPGPEGVQVGTSEKSLTWSPLQDGGSDISHY 17257  
QY 168 ----- 167  
Db 17258 VVEKRETSRLAWTVVASEVVTNSLKVTKLLEGGNEYVFRIMAVNKYGVGEPLSAPVLMKN 17317  
QY 168 ----- 167  
Db 17318 PFVLPGPKSLEVTNIAKDSMTVCWNRPDSDGSGSEIIGYIVEKRDGRGIRWIKCNKRIT 17377  
QY 168 ----- 167  
Db 17378 DLRLVTLGTEDHEYEFVRSNAENAGVGEPSATVYKACDPVKPCPPTNAHIVDTKN 17437  
QY 168 ----- 167  
Db 17438 SITLANGKPIYDGSSEILGVVVEICKADEEWQIVTPQTLRVTREISKLTEHQBKIR 17497  
QY 168 ----- 167  
Db 17498 VCAINKVGLGEATSVPGTVKPEDKLEAPDLDELSELKRGIVVRAGGSARIHIPFKGRPTP 17557  
QY 168 ----- 167  
Db 17558 EITWSREGEFTDKVOIEKGVNTQTSIDNCDRNDAGYILKLENSSGSKSAFVTVKVLD 17617  
QY 168 ----- 170  
Db 17618 TPQPPQNLAVKEYRKDSAFVLEWPEPIIDGGAKVKNYVIDKRESTRKAYANVSSKCSKTSF 17677  
QY 171 ----- 170

Db 17678 KVENLATEGAIYFRVMAENEFVGVVPVETVDVAKAAEPPSPGKVTLTDSQTSASLWE 17737  
QY 171 ----- 170  
Db 17738 KPEHDGGSRLVGVVEMQPKGTEKWSIVAESKVCNNAVVTGLSSGOEYQFRVKAYNEKGKS 17797  
QY 171 ----- 170  
Db 17798 DPRVLGVPVIKDLTIQPSLKLFPNTYSIOAGEDLKEIPVIGRPRNISMVWKGDEPLKQ 17857  
QY 171 ----- 170  
Db 17858 TTRVNVEETATSVLHIKEGKDDFGKYTVTATNSAGTATENLSVIVLEKPGVPVGRF 17917  
QY 171 ----- 170  
Db 17918 DEVSADFVVISWEPAYTGGCOISNYIVEKRDTTTTTWHMVSATVARTTIKTKLTGTE 17977  
QY 171 ----- 170  
Db 17978 YQFRIFAENRYKGSAPLDSKAVIVQYFPKPGPGTFFVTSISKDQMLVQHHEPVNDGGT 18037  
QY 171 ----- 170  
Db 18038 KLIIGYHLEQEKNSILWVKNLKTPIQDTKFTTGLDBGLEYEYFVSAENIVGIGKPSKVS 18097  
QY 171 ----- 170  
Db 18098 ECFVARPCDPPGRPEAIVITRNNVTLKWKPAYDGGSKITGYIVEKKDLPDGRWMKASF 18157  
QY 171 ----- 170  
Db 18158 TNVLETFVTVGLVEDQRYEFVRIARNAAGNFSEPSDSSGAIITARDEIDAPNASLDPKYK 18217  
QY 171 ----- 170  
Db 18218 DVIWHAGETFVLEADIRGKPIPDVVMVSKDGKELEETAARMEIKSTIQKTTLVVKDCIRT 18277  
QY 171 ----- 170  
Db 18278 DGGQYILKLSNVGTSKIPITVKVLDROPPEGLPKVTGYTAEKCYLAWNPLODGGANI 18337  
QY 171 ----- 175  
Db 18338 SHVILIEKRETSRLSWTQVSTEVOALNKKVTKLLPGNEYIFRVMVKNKYGIGEPLESGPVT 18397  
QY 176 ----- 175  
Db 18398 ACNPYKPPGPPSTPEVSATKDSMVVTWARPVDDGGTEIEGYILEKRDKEGVRTKCNKK 18457  
QY 176 ----- 175  
Db 18458 TLFDLRLRVTLGTLPEHSHYEFVRAAENAGVGEPSVFFYRACDALLYPPGPPSNPKVTD 18517  
QY 176 ----- 175  
Db 18518 SRSSVSLANSKPIYDGGAPVKGYVVEVKEAADEWTTCTPTGLQKQFTVTKLKENTEX 18577  
QY 176 ----- 175  
Db 18578 NFRICAINSEGVGEATLPGSVVVAQRIEPPPEIOLDADLRKVVVLRASATLRLFTIKGR 18637  
QY 176 ----- 175  
Db 18638 PEPEVKWEKAEGILTDRAQIEVTSSTFMTLVIDNVTRFDSGRYNTLENNSGSKTAFVNR 18697  
QY 176 ----- 175  
Db 18698 VLDSPAPVNLITREVKKDSVTLSEWPEPLIDGGAKITNYIVEKRETRKAYATINNCTK 18757  
QY 176 ----- 175

Db 18758 TTFRIENLQGCYYFPRVLASNEYGIGLPAETTEPVKVSEPLPLPGRVTLVDVTRNTATI 18817  
QY 176 ----- 175  
Db 18818 KWEPESDGGSKITGYVVEWQTKGSEKWCSTCQVKTLKLEATISGLTAGEEYVFRVAANKEK 18877  
QY 176 ----- 175  
Db 18878 GRSDPRQLGVPVIARDIEIKPSVELPFHTFNVKAREQLKIDVPFKGRPOATVNRKRGOT 18937  
QY 176 ----- 175  
Db 18938 LKETRVNVSSKTVTSLSIKEASKEDVGYELCVSNSAGSITVPITIIIVDRPPGPI 18997  
QY 176 ----- 175  
Db 18998 RIDEVSCDSITISWNPPEYDGGCQISNYIVEKKEITSTTHIVSOAVARTSIKIVRLTIG 19057  
QY 176 ----- 175  
Db 19058 SEYQFRVCAENRYGKSSYSESSAVVAEYFPSPGPGTPKVVHATKSTMLVTWQVPVNDG 19117  
QY 176 ----- 175  
Db 19118 GSRVIGYHLEYKERSIILMSKANKILLIADTQMKVSGLDEGLMYEYRVAENIAGICKSK 19177  
QY 176 ----- 175  
Db 19178 SCEPVPARPCDPGPOPEVTNITRSVSLKWSKPHYDGGAKITGYIVERRELDPGRWLKC 19237  
QY 176 ----- 175  
Db 19238 NYTNIQETYFEVTELTEDQRYEFRFARNAAADVSEPSSESTGPIIVKODVEPRVMMDVK 19297  
QY 176 ----- 175  
Db 19298 FRDIVVKAGEVLKINADIAGRPLPVISWAKDGBIEIERARTEIISTDNHLLTVKDCIR 19357  
QY 176 ----- 175  
Db 19358 RDTGVVLTKNVAGTRSVAVNKVLDKPGPPAGPLEINGLTAEKCSLSWGRPQEDGGAD 19417  
QY 176 ----- 175  
Db 19418 IDYIIVEKRETLAWTICEGELQMTSCKVKLLKGNEYIFRVTGVNKGVGCEPLESAI 19477  
QY 176 ----- 175  
Db 19478 KALDPFTVPSPPTSLEITSVTKESMTLCWSRPESDGGSEISGYIIERREKNSLRWRVVK 19537  
QY 176 ----- 175  
Db 19538 KPYVDLRVKSTGLREGCEYEVRYAENAAGLSLPSSETPLRAEDPVFLPSPSKPKIVD 19597  
QY 176 ----- 175  
Db 19598 SGKTTITIAWRKPLEDGGAPITGYVEYKSDTDWKTISIQLRGTEYTIISGLTTGAEVY 19657  
QY 176 ----- 175  
Db 19658 FRVKSVMKVGASDSDSDPOIAKEREELPFDIDSEMRKTLIVKAGASFTMTVPFRGRP 19717  
QY 176 ----- 175  
Db 19718 VPNVLSKRPDTLRTRAYVDTTDSRTSLTIENANRDSGKYTLTIQNVLSAASLTVVKV 19777  
QY 176 ----- 175  
Db 19778 LDTGPGPTNITVQDVTKESAVLSWDVPENDGGAPVKNYHIEKREASKKAWSVTNNCRN 19837  
QY 176 ----- 175  
Db 19838 SYKVTNLQEGAIYFPRVSGENEFGVIPAETKEGVKITEKPSPEKLGVTISKDSVSLT 19897

QY 176 ----- 175  
Db 19898 WLKPEHDGSRIVHYVVEALEKGQKNWKCACAVAKSTHHVVSGLRESEYFFRFAENQAG 19957  
QY 176 ----- 175  
Db 19958 LSDPRELLPVLIKQOLEPPEIDMKNFPSHTVYVRAGSNLKVDIPIISKPLPKVTLSRDG 20017  
QY 176 ----- 175  
Db 20018 VPLKATMRNTEITAENLTINLKESVTADAGRYEITAANSSGTTKAFINIVVLDRRPGPT 20077  
QY 176 ----- 175  
Db 20078 GPVWISDITEESVTLKWEPPKYDGGSQVTNYILLKRETSTAVWTEVSATVARTMMKVMKL 20137  
QY 176 ----- 175  
Db 20138 TTGEEYQFRIKAENRFGISDHIDSACVTVKLPYTTGPPSTPMVTVNTRESTITVGMHEPV 20197  
QY 176 ----- 175  
Db 20198 SNGGSVVGYHLEMKDRNSILMQKANKLVIRTHFKVTTISAGLIYEFVRVYAENAGVCK 20257  
QY 176 ----- 175  
Db 20258 PSHPSEPVLADACEPPRNVRTDISKNSVLSWQOPAFDGGSKITGYIVERRDLDPDGRW 20317  
QY 176 ----- 175  
Db 20318 TKASFTNVTETOFIISGLTQNSQYEFVRFARNAVGSI SNPSEVVGPIITCIDSYGGVIDL 20377  
QY 176 ----- 175  
Db 20378 PLEYTEVVKYRAGTSVKLRAGISGKPAPIEWYKODKELQTNALVCVNTTDLASILKD 20437  
QY 176 ----- 175  
Db 20438 ADRLNSGCYELKLRNAGSASATIRVQILDKPGPGPIEFKVTVAEKITLLARPPADGG 20497  
QY 176 ----- 175  
Db 20498 GAKITHYIIVEKRETSRVVMSVSEHLEPCIITTTKIIKNEYIFRVRVAVNKYIGIPELES 20557  
QY 176 ----- 175  
Db 20558 DSVAKNAFVTPGPPGPIPEVTKITKNSMTVWWSRPIADGGSDISGYFLEKRDKKSLGWPK 20617  
QY 176 ----- 175  
Db 20618 VLKETIRDTROKVTGLTENSQYQYRVCVNAAGQGFSEPFSEFYKKAADPIDPPGPPAKIR 20677  
QY 176 ----- 175  
Db 20678 IADSTKSSITLWGSKPVYDGGSAVTGYVVEIRQGEETTVSTRKEVTRTEYVVSNLKP 20737  
QY 176 ----- 175  
Db 20738 GVNYYFRVSAVNACGGGPIEMNEPVQAKDILEAPEIDLVALRTSVIAKAGEDVOVLIP 20797  
QY 176 ----- 175  
Db 20798 FKGRPPPTVTRKDEKNLGSADARYSIENTDSSLLTIPOVTRNDNGKYILLTIENGVEGPK 20857  
QY 176 ----- 175  
Db 20858 SSTSVKVLDTPAACQKLQVAKHVRGTVTLMDPPLIDGGSPINIVYIEKRDATKRTWSV 20917  
QY 176 ----- 175  
Db 20918 VSHKCSSTSFKLIDLSEKTPFFFRVLAENEIGICEPCTETPEPVKAAEVPAPIRDLMSKDS 20977

QY 176 ----- 175  
Db 20978 TKTSVILSWTKPFDGGSVITEYVWERKKGEOQWTHAGISKTCEIEVSQKQEVSLEPR 21037  
QY 176 ----- 175  
Db 21038 VFAKNEGLSDPVTIGTIVKELIITPEVDLSDIPGAQVTRIGHNVHLELPYKGRPKPS 21097  
QY 176 ----- 175  
Db 21098 ISWLKDLPLKESEFFVRFSKTENKITLSIKNAKEHGGKYTVILDNAVCRIAPITVITL 21157  
QY 176 ----- 175  
Db 21158 GPSPKPGPIRFDEIKADSVILSWDPEDNGGGEITCYIEKRETSQTNKMKVCSVART 21217  
QY 176 ----- 175  
Db 21218 TFKVNLVKDAEYQFRVRAENRYGVSQLVSSIIVAKHQPRIPGPGPKPVIYNWTSOGMS 21277  
QY 176 ----- 175  
Db 21278 LTWDAPYDGGSEVTGFHVEKKERNSTLMQKVNTSPISGREYRATGLVEGLDYQFRVYAE 21337  
QY 176 ----- 175  
Db 21338 NSAGLSSPSPSKFTLAVSPVDPPTGPDYIDVTRETITLKNPPLRDGSKIYGVYSIEKR 21397  
QY 176 ----- 175  
Db 21398 QGNERWVRNFTDVSECQYTVTGLSPGDRYEFRIIARNAVGTISPSPSQSGIIMTRDENV 21457  
QY 176 PPV----- 178  
|||  
Db 21458 PPIVEGPEYFDGLIIKSGESLRIKALVQGRPVPRVTFWKDGVGEIEKRMNWEITDVLGST 21517  
QY 179 ----- 178  
Db 21518 SLEVRDATDRHGYTVEAKNASGAKAEIKVKVQDTPGKVGVPIRFTNITGEKMTLWMD 21577  
QY 179 ----- 178  
Db 21578 APLNDGAPTHVIIIEKRETSRLAWALIEDKCEAQSYTAIKLINGNEYQFRVSAVNKFGV 21637  
QY 179 ---LDSD----- 182  
|||  
Db 21638 GRPLSDPVVAQIQYTVPDAPGIPEPSNITGNSITLTWARPESDGGSEIQYILERREKK 21697  
QY 183 ----- 182  
Db 21698 STRWVKVISKRPISETRFRKVTGLTEGNEYEFHYMAENAGVGPASGISRLIKCREPVNPP 21757  
QY 183 ----- 182  
Db 21758 GPPTVVKVTDTSKTVSLEWSKPVFDGGMELIGYIEMCKADLGWHKVNABACVKTRYT 21817  
QY 183 ----- 182  
Db 21818 VTDLOAGEYKFRVSAINGAGKDSCEVTCTIKAVDRLTAPELIDIDANFKQTHVVRAGAS 21877  
QY 183 ----- 182  
Db 21878 IRLFIAYQGRPTTAVWSKPDNSLSLRADIIHTTDSFSTLTVCNCRNDAGKYTLTVENNS 21937  
QY 183 ----- 182  
Db 21938 GSKSITFTVKVLDTPGPGPITFKDVTGRSATLMDAPLDDGGARTHHYVVERKESRRS 21997  
QY 183 ----- 182  
Db 21998 WQVISEKCTRQIFKVNDAEGVPYFRVSAVNEVGCEPYEMPEPIVATEQAPPRRLDV 22057  
QY 183 ----- 182

Db 22058 VDTSSSAVLAWLPKPDHGGSRITGYLLEMRQSGDFWEAGHTKQLTFTVERLVEKTEY 22117  
QY 183 ----- 182  
Db 22118 EFRVAKNDAGYSEPREAFSSVIIKEPQIEPTADLTGITNOLITCKAGSPFTIDVPISGR 22177  
QY 183 ----- 187  
-----GSFFL-----  
Db 22178 PAKVTWKLEEMRLKETDRYSIITTKDRTTLTVKDSMRGDSGRYFLTLENTAGVKTFSVT 22237  
QY 188 ----- 187  
-----  
Db 22238 VVIGRPGPVTGPIEVSSVAESCVLISWGEPKDGGTEITNYIVEKRESGTTAWQLVNSS 22297  
QY 188 ----- 187  
Db 22298 VKRTQIKVTHLTXYMEYSFRVSSNRFVSKPLESAPIIAHBFPVPPSAPTRPEVYHVSA 22357  
QY 188 ----- 187  
Db 22358 NAMSIRWEEPYHDGGSKIICYWVEKKERNITLWVKENKVPCLCNYKVTGLVEGLEVQFR 22417  
QY 188 ----- 187  
Db 22418 TYALNAAGVSKASEASRPIMAQNPVDAPGRPEVTDVTRSTVSLIWSAPAYDGGSKVVGYI 22477  
QY 188 ----- 187  
Db 22478 IERKPVSEVDGRWLKCNVTIVSDNFTVTALSEGDTYEFRLVAKNAAGVISKSESTGP 22537  
QY 188 ----- 187  
Db 22538 VTCRDEYAPKAEALDARLHGLDVTIRAGSDLVLDAAVGGKPEPKI IWTGDKELDCEKV 22597  
QY 188 ----- 187  
Db 22598 SLOYTKRATAVIAKFCDRSDSGKYTLTVKNASGTVKAVSMVKVLDSPGCKLTVSRVTQ 22657  
QY 188 ----- 187  
Db 22658 EKTLAWSLPQEDGGAETHYIVERRETSRLNWVIVEGECPTLSYVVVTRLIKNEYIFRV 22717  
QY 188 ----- 187  
Db 22718 RAVNKYGPVVESEPIVARNSTIPSPGIPPEVGTGKEHIIQWTKPESDGGNEISNY 22777  
QY 188 ----- 187  
Db 22778 LVDKREKKSRLWTRVKNKYVYDTRLKVTSLMEGCDYQFRVAVNAAGNSEPSEASNFIS 22837  
QY 188 ----- 187  
Db 22838 CREPSYTPGPPSAPRVVDTTKHSISLAWTKPMYDGGTDIVGYVLEMQEKTDQWYRVHTN 22897  
QY 188 ----- 187  
Db 22898 ATIRNTEFTVPLDKMGOKYSFRVAANVVKGMSEYSEIAETEPVERIEIPDLELADLKK 22957  
QY 188 ----- 195  
-----YSKLTVDK-----  
|||  
Db 22958 TVTIRAGASRLMVSVSGRPPPVITWSKOGIDLASRAIIDTTESYSLLIIVDKVNRDAGK 23017  
QY 196 ----- 195  
Db 23018 YTIEAENQSGKKSATVLVKYDTPGCPSPVKVEVSRDSVTITWEITIDGGAPVNNYIV 23077  
QY 196 ----- 195  
Db 23078 EKREAAAMRAFKTVTKCSKTLYRISGLVEGTYMYFRVLPENIYIGIPECTSDAVLYSEV 23137  
QY 196 ----- 195

Db 23138 PLVPKLEVVDTKSTVTLANEKPLYDGSRLTGYVLEACKAGTERMMKVVTLPKTVLEH 23197  
QY 196 ----- 195  
Db 23198 TVTSLNEGQYLFRIRAOKEGVSEPREIVTAVTVQDLRVLTIDLTSTMPQKTIHVPAGR 23257  
QY 196 ----- 195  
Db 23258 PVELVIAGRPAPPAASFAGSKLRESERVTVETHKVAKLTIRETTIRDTGEVTELEK 23317  
QY 196 ----- 195  
Db 23318 NVTGTTSETIKVILDKPGPTGPIKIDEIDATISISWPELDPGAPLSGYVVEQRDA 23377  
QY 196 ----- 195  
Db 23378 HRPGLPVSESVTRSTFKETRLTEGNEYVFRVAATNRFGISYLOSEVIECRSSIRIPCP 23437  
QY 196 ----- 195  
Db 23438 PETLIQFVSRDGMTLTWYPPEDDGGSQVTGYIVERKEVRADRWVRVNVKVPVTMTYRST 23497  
QY 196 ----- 195  
Db 23498 GLTEGLEVHRVTAI NARGSGKPSRSPKPIVAMPPIAPPKQPNRPVTDTRTSVSLAWS 23557  
QY 196 ----- 195  
Db 23558 VPDEGGSKVTGLIEMQKVDOHEWTKNTTPTKIREYTLTHLPOGAEYRFRVLACNAGG 23617  
QY 196 ----- 201  
Db 23618 PGEPAEVCTVKVTEMLEYPDYELDERYQEGIFVRGGVIRLTIPKGRPFPICKWTKREG 23677  
QY 202 ----- 201  
Db 23678 QDISKRATATSETHVELVIKEADRGDSGTVDLVLENKCGKAVYIKVRVIGSPNSPEGP 23737  
QY 202 ----- 201  
Db 23738 LEYDDIQRVSRVSRPPADGGADILGYILERREVPRKAANYTIDSRVRGTSLVVKGLE 23797  
QY 202 ----- 201  
Db 23798 NVEYHFRVAENQFGISKPLKSEEPVTKPLNPPPEPNPEVLDVTKSSVLSWSRPK 23857  
QY 202 ----- 201  
Db 23858 DDGSRVTGYIETKSTDKVVRHNKQITTTMTYTVTGLVPDAEYQFRIIAQNDVGLSE 23917  
QY 202 ----- 201  
Db 23918 TSPASEPVVCKDPFDKPSOPGGELEILSISKDSVTLOWEKPCEGCKEILGYWVEYRQSGD 23977  
QY 202 ----- 201  
Db 23978 SAWKSKNERIKDKOFTIGGLLEATEYEFVRPAENETGLSRPRRTAMSIKLTSGEAPG 24037  
QY 202 ----- 201  
Db 24038 IRKENKDVTKLGEAAQLSCQIVGRPLPDIKWYRFGKELIOSRKYKMSSDGRTHLTVMY 24097  
QY 202 ----- 201  
Db 24098 EEQDEGVVTCIATNEVEGETSSKLLQATPQPHPGYPLKEKYGAVGSTLRLHVMYIG 24157  
QY 202 ----- 204  
Db 24158 RVPAMTFHGQKLLQNSENITIENTEYTHLVKMNQVQKTHAGKYKVQLSNVFGTVDAI 24217  
QY 205 ----- 204  
Db 24218 LDVEIQDKPDKPTGPVIVIEALLKNSAVISWKPADDGGSWITNYYVEKCEAKEGEWOLV 24277

QY 205 ----- 204  
Db 24278 SSAISVTTTCRIVNLTENAGYFRVSAQNTFGISDPLEVSSVVIKSPPEKPGAPGKPTIT 24337  
QY 205 ----- 208  
Db 24338 AVTKDSCVAVMKPPASDGGAKIRNYYLEKREKKQNKWISVTTTEIRETVFSVKNLIEGLE 24397  
QY 209 ----- 208  
Db 24398 YEFVRKCNLGESESEISEPITPKSDVPTQAPHKBEELNMLNVRYSNATLVCKVTCGH 24457  
QY 209 ----- 208  
Db 24458 PKPIVKNYRQKEIADGLKYRIQEFKGGYHOLIIASVTDDBDATVYQVRATNOGGSVST 24517  
QY 209 ----- 208  
Db 24518 ASLEVEVPAKIHLPKLTLEGMGAVHALRGVSVIKIPFSGKPDVITWQKQDLIDNNGHY 24577  
QY 209 ----- 208  
Db 24578 QVIVTRSFVSLVPNGVERKDGAFYVVCANKNFGIDQKTVELDVADVPDPGRGVKVDVS 24637  
QY 209 ----- 208  
Db 24638 RDSVNLWTEPASDGGSKITNYIVEKCATTAERWLRVQARETRYVINLFGKTSYQVRV 24697  
QY 209 ----- 208  
Db 24698 IAKNFGLSKPSEPSEPTITKEDKTRAMNYDEEVDETVSMTKASHSSTKELYEKYMA 24757  
QY 209 ----- 212  
Db 24758 EDLGRGEFIVHRCVETSSKTYMAKFKVKGTDQVLVKKEISILNIAHRNHLHESF 24817  
QY 213 ----- 216  
Db 24818 ESMEELVMIFEFISOLDIFERINTSAFELNBEREIVSVHVQVCEALQFLHSHNIGHDIRP 24877  
QY 217 ----- 216  
Db 24878 ENIIYOTRRSSTIKIIEFGARQLKPGDNFRLLFTAPEYAYEVEHQHDVSTATDMWSLG 24937  
QY 217 ----- 218  
Db 24938 TLVVYVLLSGINPFLAETNQIENIMNAEYTFDEEAFKEISIEAMDFVDRLLVKKRGRM 24997  
QY 219 ----- 221  
Db 24998 TASEALQHPWLKQKTIERYSTKVIRTLKHRRYYHTLTKKDLNWNVSAARISCGGAIRSKG 25057  
QY 222 LSLS ----- 225  
Db 25058 VSAKVKVASIEIGPVSGQIMHAVGEGGHVKYCKIENYDQSTQVTVFGVRLNSEK 25117  
QY 226 ----- 225  
Db 25118 YEITYEDGVAILYVKDITKLDGTYRCKVNDYGEDSSYAEFLVKGVREVYDYCRRTMK 25177  
QY 226 ----- 225  
Db 25178 KIKRRTDMRLLEPPEFTPLYNKTA VGENVREGVITVHPEPHVTWYKSGOKIKPGD 25237  
QY 226 ----- 225  
Db 25238 NDKYTFESDKGLYQLTINSVTDDDAEYTVARNKYGEDSKAKLTVTLHPPPTDSTLR 25297  
QY 226 ----- 225  
Db 25298 PMFKRLLANABEQSGSVCFEIRVSGIPPTLKWEKQPLSLGNIEIIEHGLDYALH 25357







GenCore version 5.1.4\_p5\_4578  
Copyright (c) 1993-2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: April 21, 2003, 10:42:21 ; Search time 31.0588 Seconds  
(without alignments)  
1132.630 Million cell updates/sec

Title: 2LINK7LINK7

Perfect score: 1430

Sequence: 1 MDKTHTCPCPAPPELLGGPS.....AFDWLKAIFYDKVAEKLKEAF 264

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.0

Searched: 908470 seqs, 133250620 residues

Total number of hits satisfying chosen parameters: 908470

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

A\_Geneseq\_101002:\*

- 1: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1980.DAT:\*
- 2: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1981.DAT:\*
- 3: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1982.DAT:\*
- 4: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1983.DAT:\*
- 5: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1984.DAT:\*
- 6: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1985.DAT:\*
- 7: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1986.DAT:\*
- 8: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1987.DAT:\*
- 9: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1988.DAT:\*
- 10: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1989.DAT:\*
- 11: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1990.DAT:\*
- 12: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1991.DAT:\*
- 13: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1992.DAT:\*
- 14: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1993.DAT:\*
- 15: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1994.DAT:\*
- 16: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1995.DAT:\*
- 17: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1996.DAT:\*
- 18: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1997.DAT:\*
- 19: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1998.DAT:\*
- 20: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1999.DAT:\*
- 21: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA2000.DAT:\*
- 22: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA2001.DAT:\*
- 23: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA2002.DAT:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Match	Length	ID	Description
1	1260	88.1	729	22 AAM52158	Humanised HMFG-1 h
2	1260	88.1	739	22 AAM52161	Humanised HMFG-1 h
3	1257	87.9	731	22 AAM52156	Humanised HMFG-1 h
4	1257	87.9	741	22 AAM52159	Humanised HMFG-1 h
5	1253	87.6	689	21 AAY96780	Ang-2-FD-Fc-FD fus
6	1251	87.5	652	19 AAW48650	Heavy chain of hma
7	1250	87.4	423	21 AAE28693	Fc-huAcp-1 (114-28
8	1250	87.4	441	21 AAE28692	Fc-huAcp-1 (95-281
9	1250	87.4	730	22 AAM52157	Humanised HMFG-1 h
10	1250	87.4	740	22 AAM52160	Humanised HMFG-1 h

11	1249	87.3	633	21 AAY84965	Amino acid sequenc
12	1247	87.2	426	21 AAB28694	Fc-huAcp-1 (120-29
13	1247	87.2	448	21 AAB28694	Fc-huAcp-1 (99-291
14	1246	87.1	247	21 AAB16958	Fc-TMP protein seq
15	1246	87.1	247	23 AAB73411	Fc-TPO mimetic pep
16	1246	87.1	248	21 AAB71953	Fc-IL-1 antagonist
17	1246	87.1	248	23 AAB73421	Fc-interleukin 1 (
18	1246	87.1	268	21 AAB16959	Fc-TMP-TMP protein
19	1246	87.1	268	23 AAB73412	Fc-TMP-TMP amino a
20	1246	87.1	269	21 AAY96531	Human IgG1 Fc TMP
21	1246	87.1	374	19 AAW83963	Recombinant human
22	1246	87.1	374	19 AAW49075	Recombinant human
23	1246	87.1	401	22 AAY72922	Human met-Fc (lack
24	1246	87.1	401	22 AAB80904	Human metFcDeltaC-
25	1246	87.1	689	21 AAY96779	Ang-1-FD-Fc-FD fus
26	1245	87.1	252	21 AAB17955	Fc-VEGF antagonist
27	1245	87.1	252	23 AAB73423	Fc-VEGF antagonist
28	1245	87.1	651	18 AAW26649	Chimeric receptor
29	1245	87.1	692	18 AAW26650	Chimeric receptor
30	1241	86.8	379	19 AAW83962	Recombinant human
31	1241	86.8	379	19 AAW49073	Recombinant human
32	1241	86.8	379	19 AAW49074	Recombinant human
33	1240	86.7	389	22 AAW50247	Human IgG1 Fc-inte
34	1240	86.7	581	22 AAB81972	Ganglioside GD2 sp
35	1240	86.7	582	22 AAB81987	Ganglioside GD3 sp
36	1240	86.7	582	22 AAB81991	Ganglioside GD3 sp
37	1240	86.7	583	22 AAB83156	Ganglioside GM2 an
38	1239	86.6	248	21 AAB17951	Fc-TNF-alpha inhib
39	1239	86.6	248	23 AAB73419	Fc-TNF-alpha inhib
40	1239	86.6	253	21 AAB16964	Fc-EMP protein seq
41	1239	86.6	253	23 AAB73415	Fc-EPO mimetic pep
42	1239	86.6	277	21 AAB16967	Fc-EMP-EMP protein
43	1239	86.6	277	23 AAB73418	Fc-EMP-EMP nucleic
44	1239	86.6	660	20 AAX13463	Ectromelia A39R se
45	1239	86.6	660	21 AAB28523	Ectromelia A39R se

#### ALIGNMENTS

RESULT 1

AAM52158

ID AAM52158 standard; Protein; 729 AA.

AC AAM52158;

XX AAM52158;

DT 05-FEB-2002 (first entry)

XX Humanised HMFG-1 heavy chain/DNase I fusion protein 3.

DE DE

XX Humanised monoclonal antibody; polymorphic epithelial mucin; PEM1;

KW cytotoxic; endonuclease; DNase I; human; cytostatic; cancer; apoptosis.

XX Homo sapiens.

OS Synthetic.

XX WO200174905-A1.

PN 11-OCT-2001.

XX 26-MAR-2001; 2001WO-GB01324.

XX 03-APR-2000; 2000GB-0008049.

PR 02-OCT-2000; 2000US-237159P.

XX (ANTI-) ANTISOMA RES LTD.

XX Young RJ;

XX WPI; 2001-662969/76.

XX Novel compound used to treat cancer has target cell-specific portion

XX comprising humanised monoclonal antibody having specificity for

PT

PT polymorphic epithelial mucin, and cytotoxic portion having  
XX endonucleolytic activity  
PS Claim 20; Figure 9; 176pp; English.  
XX The invention relates to a compound which comprises a target  
CC cell-specific portion, comprising an humanised monoclonal antibody,  
CC having specificity for polymorphic epithelial mucin (PEM) or its antigen  
CC binding fragment and a cytotoxic portion having endonucleolytic activity,  
CC exemplified by AAM52154-AAM52168 and encoded by ABA02682-ABA02728. The  
CC compound has cytostatic activity useful for treating cancer and acting as  
CC a potential inducer of apoptosis.  
XX Sequence 729 AA;  
SQ Query Match 88.1%; Score 1260; DB 22; Length 729;  
Best Local Similarity 50.5%; Pred. No. 3.7e-33;  
Matches 241; Conservative 6; Mismatches 15; Indels 215; Gaps 5;  
QY 2 DKHTCCPCPAPPELLGGPSVFLFPPKPKDTLMISRTPEVTCVVDVSHEDPEVKFNWYVD 61  
Db 241 DKHTCCPCPAPPELLGGPSVFLFPPKPKDTLMISRTPEVTCVVDVSHEDPEVKFNWYVD 300  
QY 62 GVEVHNKTKPREEOYNSTYRVSVLTCLVKGFPSPDIAVEWESNGOPENNYKTTTPVLD 121  
Db 301 GVEVHNKTKPREEOYNSTYRVSVLTCLVKGFPSPDIAVEWESNGOPENNYKTTTPVLD 360  
QY 122 GOPREPQVYTLPPSRDELTKNQVSLTCLVKGFYPSDIAVEWESNGOPENNYKTTTPVLD 181  
Db 361 GOPREPQVYTLPPSRDELTKNQVSLTCLVKGFYPSDIAVEWESNGOPENNYKTTTPVLD 420  
QY 182 DGSFFLYSKLTVDKSRWQGNVFCSCVMHEALHNHYTQKSLSLSPGKDWL 232  
Db 421 DGSFFLYSKLTVDKSRWQGNVFCSCVMHEALHNHYTQKSLSLSPGKDWL 480  
QY 233 -----KAFYDKV-----AEKLKEA 263  
Db 661 FQWLIPDSADTTATPTHCAIDRIYVAGMLLRGAVVPDSALPFNFQAAYGLSDQLAQA 717  
RESULT 2  
AAM52161  
ID AAM52161 standard; Protein; 739 AA.  
XX AAM52161;  
AC AAM52161;  
XX 05-FEB-2002 (first entry)  
XX Humanised HMFG-1 heavy chain/DNase I fusion protein 6.  
DE Humanised monoclonal antibody; polymorphic epithelial mucin; PEM;  
KW cytotoxic; endonuclease; DNase I; human; cytostatic; cancer; apoptosis.  
XX Homo sapiens.  
OS Synthetic.  
XX WO200174905-A1.  
PN 11-OCT-2001.  
XX 26-MAR-2001; 2001WO-GB01324.

XX 03-APR-2000; 2000GB-0008049.  
PR 02-OCT-2000; 2000US-237159P.  
XX (ANTI-) ANTISOMA RES LTD.  
XX Young RJ;  
PI WPI; 2001-662969/76.  
DR Novel compound used to treat cancer has target cell-specific portion  
PT comprising humanised monoclonal antibody having specificity for  
PT polymorphic epithelial mucin, and cytotoxic portion having  
PT endonucleolytic activity  
XX Claim 20; Figure 12; 176pp; English.  
XX The invention relates to a compound which comprises a target  
CC cell-specific portion, comprising an humanised monoclonal antibody,  
CC having specificity for polymorphic epithelial mucin (PEM) or its antigen  
CC binding fragment and a cytotoxic portion having endonucleolytic activity,  
CC exemplified by AAM52154-AAM52168 and encoded by ABA02682-ABA02728. The  
CC compound has cytostatic activity useful for treating cancer and acting as  
CC a potential inducer of apoptosis.  
XX Sequence 739 AA;  
SQ Query Match 88.1%; Score 1260; DB 22; Length 739;  
Best Local Similarity 50.5%; Pred. No. 3.7e-33;  
Matches 241; Conservative 6; Mismatches 15; Indels 215; Gaps 5;  
QY 2 DKHTCCPCPAPPELLGGPSVFLFPPKPKDTLMISRTPEVTCVVDVSHEDPEVKFNWYVD 61  
Db 241 DKHTCCPCPAPPELLGGPSVFLFPPKPKDTLMISRTPEVTCVVDVSHEDPEVKFNWYVD 300  
QY 62 GVEVHNKTKPREEOYNSTYRVSVLTCLVKGFPSPDIAVEWESNGOPENNYKTTTPVLD 121  
Db 301 GVEVHNKTKPREEOYNSTYRVSVLTCLVKGFPSPDIAVEWESNGOPENNYKTTTPVLD 360  
QY 122 GOPREPQVYTLPPSRDELTKNQVSLTCLVKGFYPSDIAVEWESNGOPENNYKTTTPVLD 181  
Db 361 GOPREPQVYTLPPSRDELTKNQVSLTCLVKGFYPSDIAVEWESNGOPENNYKTTTPVLD 420  
QY 182 DGSFFLYSKLTVDKSRWQGNVFCSCVMHEALHNHYTQKSLSLSPGKDWL 232  
Db 421 DGSFFLYSKLTVDKSRWQGNVFCSCVMHEALHNHYTQKSLSLSPGKDWL 480  
QY 233 -----KAFYDKV-----AEKLKEA 263  
Db 481 GETKMSNATLVSYIVQILSRDIALVQEVDRSHLTAVGKLLDNLDNODAPDTYHYVVEPL 540  
QY 233 -----KAFYDKV-----AEKLKEA 263  
Db 541 GRNSYKERYLFVYRPDQVSAVDYDDGCEPCGNDTFNREPAIVRFFSRTFVREFAIV 600  
QY 233 -----AFYD---KVAEK-----LKEA 245  
Db 601 PLHAAPGDAVEIDALDYVDVQEKWGLDVLMGDFNAGGCVYVRPSQWSSIRLWTSPT 660  
QY 246 FQWL-----KAFYDKV-----AEKLKEA 263  
Db 661 FQWLIPDSADTTATPTHCAIDRIYVAGMLLRGAVVPDSALPFNFQAAYGLSDQLAQA 717  
RESULT 3  
AAM52156  
ID AAM52156 standard; Protein; 731 AA.  
XX AAM52156;  
AC AAM52156;  
XX 05-FEB-2002 (first entry)  
XX Humanised HMFG-1 heavy chain/DNase I fusion protein 1.

XX Humanised monoclonal antibody; polymorphic epithelial mucin; PEM1;  
KW cytotoxic; endonuclease; DNase I; human; cytostatic; cancer; apoptosis.  
XX Homo sapiens.  
OS Synthetic.  
XX WO200174905-A1.  
XX 11-OCT-2001.  
XX 26-MAR-2001; 2001WO-GB01324.  
XX 03-APR-2000; 2000GB-0008049.  
XX 02-OCT-2000; 2000US-237159P.  
XX (ANTI-) ANTISOMA RES LTD.  
XX Young RJ;  
XX WPI; 2001-662969/76.  
XX Novel compound used to treat cancer has target cell-specific portion  
PT comprising humanised monoclonal antibody having specificity for  
PT polymorphic epithelial mucin, and cytotoxic portion having  
PT endonucleolytic activity  
XX Claim 20; Figure 7; 176pp; English.  
XX The invention relates to a compound which comprises a target  
CC cell-specific portion, comprising an humanised monoclonal antibody,  
CC having specificity for polymorphic epithelial mucin (PEM) or its antigen  
CC binding fragment and a cytotoxic portion having endonucleolytic activity,  
CC exemplified by AAM52154-AM52168 and encoded by ABA02682-ABA02728. The  
CC compound has cytostatic activity useful for treating cancer and acting as  
CC a potential inducer of apoptosis.  
XX Sequence 731 AA;  
SQ  
Query Match 87.9%; Score 1257; DB 22; Length 731;  
Best Local Similarity 50.1%; Pred. No. 4.6e-33;  
Matches 240; Conservative 6; Mismatches 16; Indels 217; Gaps 4;  
QY 2 DKHTCPCPAPELLGGPSVFLPFPKPKDTLMISRTPEVTCVVDVSHEDPEVKFNMYVD 61  
Db 241 DKHTCPCPAPELLGGPSVFLPFPKPKDTLMISRTPEVTCVVDVSHEDPEVKFNMYVD 300  
QY 62 GVEVHNAKTPREEQYNSTYRVVSVLTVLHQDWLNGKEYKCKVSNKALPAPIEKTISKAK 121  
Db 301 GVEVHNAKTPREEQYNSTYRVVSVLTVLHQDWLNGKEYKCKVSNKALPAPIEKTISKAK 360  
QY 122 GQPREPOVYTLPPSRDELTKNQVSLTCLVKGFYPSDIAVWESNGQPENNYKTTPPVLD 181  
Db 361 GQPREPOVYTLPPSRDELTKNQVSLTCLVKGFYPSDIAVWESNGQPENNYKTTPPVLD 420  
QY 182 DGSFFLYSKLTVDKSRWQOQGVNFSCSVHREALNHHYTKLSLSLSPGK----- 228  
Db 421 DGSFFLYSKLTVDKSRWQOQGVNFSCSVHREALNHHYTKLSLSLSPGK----- 480  
QY 229 -----DKAFYDKVAEK----- 228  
Db 481 TFGETKMSNATLSYIVQILSRDYIALVQEVDRSHLTVAGKLLDNLNQADPTYYHYVSE 540  
QY 229 ----- 228  
Db 541 PLGRNSYKERYLFYRPDPQVSAVDSSYYDDGCEPCGNDTFNREPAIVRFFSRFTEVREFA 600  
QY 229 -----DWLKAFFDKVAEK-----LK 243  
Db 601 IVPLHAAPGDAVAEIDALVDYLDVQEKWGLDVMGLMGDFNAGCSYVRPQWSIRLWTS 660  
QY 244 EAFDWL-----KAFYDKV-----AEKLKEA 263  
| | | | |  
| | | | |

Db 661 PTFOWLIPDSADTTATPTHTCAYDRIVVAGMLLRGVAVPDSALPFFOAAVGLSDQLAQA 719  
RESULT 4  
AAM52159  
ID AAM52159 standard; Protein; 741 AA.  
XX AAM52159;  
XX AC AAM52159;  
XX DT 05-FEB-2002 (first entry)  
XX XX Humanised HMFG-1 heavy chain/DNase I fusion protein 4.  
XX DE Humanised monoclonal antibody; polymorphic epithelial mucin; PEM1;  
XX KW cytotoxic; endonuclease; DNase I; human; cytostatic; cancer; apoptosis.  
XX OS Homo sapiens.  
OS Synthetic.  
XX WO200174905-A1.  
XX PN 11-OCT-2001.  
XX PD 26-MAR-2001; 2001WO-GB01324.  
XX PF 03-APR-2000; 2000GB-0008049.  
XX PR 02-OCT-2000; 2000US-237159P.  
XX XX (ANTI-) ANTISOMA RES LTD.  
XX PA Young RJ;  
XX PI WPI; 2001-662969/76.  
XX DR Novel compound used to treat cancer has target cell-specific portion  
PT comprising humanised monoclonal antibody having specificity for  
PT polymorphic epithelial mucin, and cytotoxic portion having  
PT endonucleolytic activity  
XX Claim 20; Figure 10; 176pp; English.  
XX The invention relates to a compound which comprises a target  
CC cell-specific portion, comprising an humanised monoclonal antibody,  
CC having specificity for polymorphic epithelial mucin (PEM) or its antigen  
CC binding fragment and a cytotoxic portion having endonucleolytic activity,  
CC exemplified by AAM52154-AM52168 and encoded by ABA02682-ABA02728. The  
CC compound has cytostatic activity useful for treating cancer and acting as  
CC a potential inducer of apoptosis.  
XX Sequence 741 AA;  
SQ  
Query Match 87.9%; Score 1257; DB 22; Length 741;  
Best Local Similarity 50.1%; Pred. No. 4.7e-33;  
Matches 240; Conservative 6; Mismatches 16; Indels 217; Gaps 4;  
QY 2 DKHTCPCPAPELLGGPSVFLPFPKPKDTLMISRTPEVTCVVDVSHEDPEVKFNMYVD 61  
Db 241 DKHTCPCPAPELLGGPSVFLPFPKPKDTLMISRTPEVTCVVDVSHEDPEVKFNMYVD 300  
QY 62 GVEVHNAKTPREEQYNSTYRVVSVLTVLHQDWLNGKEYKCKVSNKALPAPIEKTISKAK 121  
Db 301 GVEVHNAKTPREEQYNSTYRVVSVLTVLHQDWLNGKEYKCKVSNKALPAPIEKTISKAK 360  
QY 122 GQPREPOVYTLPPSRDELTKNQVSLTCLVKGFYPSDIAVWESNGQPENNYKTTPPVLD 181  
Db 361 GQPREPOVYTLPPSRDELTKNQVSLTCLVKGFYPSDIAVWESNGQPENNYKTTPPVLD 420  
QY 182 DGSFFLYSKLTVDKSRWQOQGVNFSCSVHREALNHHYTKLSLSLSPGK----- 228  
Db 421 DGSFFLYSKLTVDKSRWQOQGVNFSCSVHREALNHHYTKLSLSLSPGK----- 480  
QY 229 -----AEKLKEA 263  
| | | | |  
| | | | |

Db 481 TFGETKMSNATLYSYVQILSRDYDIALVOEVRDLSHTAVGKLLDNLNQADPTDTHYVWSE 540  
QY 229 ----- 228  
Db 541 PLGRNSYKERYLFVYRPDQVSAVDSYVDDGCEPCGNDTFNREPAIVRFFSRFTEVREFA 600  
QY 229 -----DWLKAFYDKVAEK-----LK 243  
Db 601 IVPFLHAAPGDAVAEIDALYDVLDOVKWGLDVMMLGDFNAGCSYVRPSSIRLWTS 660  
QY 244 EAFDWL-----KAFYDKV-----AEKLKEA 263  
Db 661 PTFQWLLPDSADTATPTTCANDRIVVAGMLLRGAVVPDSALPFFNQAAAYGLSDQLAQA 719  
RESULT 5  
AA96780  
ID AAY96780 standard; Protein; 689 AA.  
XX  
AC AAY96780;  
XX  
DT 26-SEP-2000 (first entry)  
XX  
XX Ang-2-FD-FC fusion protein.  
DE  
DE Angiopoietin-2; Tie-2 receptor; ligand; antagonist; fibrinogen domain;  
KW dimer; Fc domain; fusion protein; anti-proliferative; cytostatic;  
KW leukaemia; haematopoietic factor.  
XX  
OS Chimeric - Homo sapiens.  
XX Chimeric - Synthetic.  
FH Key Location/Qualifiers  
FT Peptide 1..15  
FT /label= Trypsin\_signal\_sequence  
FT Domain 16..230  
FT /label= Ang-2\_fibrinogen-like\_domain  
FT Peptide 231..234  
FT /label= Bridging\_peptide  
FT Region 235..466  
FT /label= Human\_IgG1\_Fc\_region  
FT Peptide 467..474  
FT /label= Bridging\_peptide  
FT Domain 475..689  
FT /label= Ang-2\_fibrinogen-like\_domain  
XX  
PN WO200037642-A1.  
XX  
XX 29-JUN-2000.  
PD  
XX  
PF 23-DEC-1999; 98WO-US30900.  
PR  
XX 23-DEC-1998; 98US-0113387.  
PR  
XX (REGE-) REGENERON PHARM INC.  
PA  
XX Davis SJ, Gale NW, Yancopoulos GD, Stahl N;  
PI WPI; 2000-442670/38.  
XX N-PSDB; AAA51344.  
DR  
XX Polynucleotide encoding a fusion polypeptide, useful for promoting  
PT differential function and influencing phenotype, comprises two subunits  
PT containing at least one copy of the receptor binding domain of a ligand  
XX  
XX Example 1; Fig 4A-E; 97pp; English.  
PS  
XX Angiopoietin-1 (Ang-1) is a Tie-2 receptor ligand and is an agonist for  
CC Tie-2, whereas angiopoietin-2 (Ang-2) is a naturally occurring antagonist  
CC of the Tie-2 receptor. The fibrinogen domains (FD) of Ang-1 and Ang-2 are  
CC the receptor-binding domains and dimerized versions, of e.g. Ang-1-FD-FC  
CC (Ang-1 fibrinogen domain fused to an Fc domain), can bind to the Tie-2  
CC receptor with much higher affinity than monomeric Ang-1-FD (dimerization

CC occurs between the Fc components of adjacent molecules). However,  
CC Ang-1-FD-FC is not able to induce phosphorylation (activate) the Tie-2  
CC receptor on endothelial cells unless it is further clustered with goat  
CC anti-human Fc antibodies. The novel fusion proteins, mutant versions of  
CC Ang-1-FD and Ang-2-FD, were designed that were intrinsically more highly  
CC clustered. Tie-2 agonist fusion proteins may be used as haematopoietic  
CC factors. Tie-2 receptor antagonist fusion proteins may be used to  
CC diagnose or treat, e.g. myeloproliferative or other proliferative  
CC disorders of blood forming organs, e.g. thrombocythemias, polycythemias  
CC and leukemias. Production of homogeneous forms of clustered ligands is  
CC broadly applicable to improve the affinity and/or increase the activity  
CC of a ligand as compared to the native form of the ligand. Ephrin fusion  
CC proteins have also been constructed, which may be useful for treating  
CC neurological disorders. The ephrin fusion proteins are preferably capable  
CC of binding to Elk receptor and are especially Efl-6 antagonists.  
XX  
SQ Sequence 689 AA;  
Query Match 87.6%; Score 1253; DB 21; Length 689;  
Best Local Similarity 55.1%; Pred. No. 5.4e-33;  
Matches 238; Conservative 7; Mismatches 9; Indels 178; Gaps 4;  
QY 2 DKTHTCPPCAPPELLGGPSVFLFPPPKDITLMISRTPEVTCVVVDVSHEDPEVKENWYVD 61  
Db 240 DKTHTCPPCAPPELLGGPSVFLFPPPKDITLMISRTPEVTCVVVDVSHEDPEVKFNWYVD 299  
QY 62 GVEVHNATKPREQYNSTYRVVSVLTVLHQDWLNGKEYKCKVSNKALPAPIETISKAK 121  
Db 300 GVEVHNATKPREQYNSTYRVVSVLTVLHQDWLNGKEYKCKVSNKALPAPIETISKAK 359  
QY 122 GQPREPQVYTLPPSRDELTKNQVSLTCLVKGFYPSDIAVEWESNGQPENNYKTTPPVLD 181  
Db 360 GQPREPQVYTLPPSRDELTKNQVSLTCLVKGFYPSDIAVEWESNGQPENNYKTTPPVLD 419  
QY 182 DGSFELYSKLVYDKSRWQGNVFCSCVMHEALHNNHYTQKLSLSPGK----- 228  
Db 420 DGSFELYSKLVYDKSRWQGNVFCSCVMHEALHNNHYTQKLSLSPGKGGGSGAPRDCAE 479  
QY 229 -----DWLKAFYDKVA-----EKLKEAF 246  
Db 480 VFKSGHTTNGIYTLTFPNSTEEIKAYCDMEAGGGGWTIIQRREDGSDVDFQRTWKEYKVG 539  
QY 247 -----DW----- 248  
Db 540 GNPGEYWLGNFVSQLTNQORYVLKIHLKDWEKNGEAYSLYEHFVLSSELNRYRIHLKGL 599  
QY 249 ----- 248  
Db 600 TGTAGKISSISQPGNDFSTKGDNDKICKCSQMLTGGWFDACGSPSNLNGMYYPQRONT 659  
QY 249 -----LKAFYDK 255  
Db 660 NKENGIKWYWK 671  
RESULT 6  
AAW48650  
ID AAW48650 standard; Protein; 652 AA.  
XX  
AC AAW48650;  
XX  
XX 04-AUG-1998 (first entry)  
XX Heavy chain of hmAb425 fused to TNF alpha.  
XX Antibody-cytokine fusion protein; tricistronic vector; chimeric;  
XX TNF alpha; IL-2; IRES; internal ribosome entry site.  
OS Synthetic.  
OS Homo sapiens.  
XX  
FH Key Location/Qualifiers  
FT Region 1..494

FT /note= "Heavy chain of human mAb 425"  
FT 495..652  
FT /note= "TNF alpha"

XX WO9811241-A1.

XX 19-MAR-1998.

XX 02-SEP-1997; 97WO-EP04765.

XX 30-SEP-1996; 96EP-0115635.

PR 16-SEP-1996; 96EP-0114820.

XX (MERE ) MERCK PATENT GMBH.

XX Bruenner W, Burge C, Dunker R, Hauser H, Mielke C;

PI Rieke E, Von Hoegen I, Welge I;

XX WPI: 1998-207400/18.

DR N-PSDB; AAV18096.

XX Oligo:cistronic expression vector - useful for production of, e.g.  
PT MAb425/TNF- $\alpha$  or MAb425/IL-2 antibody fusion protein  
PT Disclosure; Fig 15; 89pp; English.

XX The present sequence represents a fusion protein comprising of TNF  
CC alpha fused to the C-terminus of the heavy chain of the human  
CC monoclonal antibody 425 (hmb425). The hmb425 has specificity for  
CC the human EGF receptor. The invention claims for a new pmCLDHAP  
CC tristicronic vector (AAV18096) for the expression of an  
CC antibody-cytokine fusion protein, hmb425-TNF alpha. The TNF alpha  
CC sequence can be substituted by the IL-2 sequence. The vector also  
CC contains a strong promoter/enhancer unit, a selection marker gene and at  
CC least two poliovirus derived internal ribosomal entry site (IRES)  
CC sequences. The vector can be expressed in mammalian host cells for the  
CC production of heteromeric fusion proteins. This expression system is  
CC claimed to produce the heteromeric proteins in high yields.

XX Sequence 652 AA;

Query Match 87.5%; Score 1251; DB 19; Length 652;  
Best Local Similarity 65.4%; Pred. No. 5.6e-33;  
Matches 238; Conservative 8; Mismatches 13; Indels 105; Gaps 5;

QY 2 DKHTCCPCPAPELLGGPSVFLPFPKPKDTLMISRTPEVTCVVVDVSHEDPEVKFNMYVD 61

Db 268 DKHTCCPCPAPELLGGPSVFLPFPKPKDTLMISRTPEVTCVVVDVSHEDPEVKFNMYVD 327

QY 62 GVEVHNAKTPREEQYNSTYRVVSVLTVLHQDWLNGKEYKCKVSNKALPAPIETISKAK 121

Db 328 GVEVHNAKTPREEQYNSTYRVVSVLTVLHQDWLNGKEYKCKVSNKALPAPIETISKAK 387

QY 122 GQPREPOVYLLPPSRDELTKNQVSLTCLVKGFYPSDIAVEWESNGQPENNYKTTTPPVLD 181

Db 388 GQPREPOVYLLPPSRDELTKNQVSLTCLVKGFYPSDIAVEWESNGQPENNYKTTTPPVLD 447

QY 182 DGSFFLYSKLTVDKSRWQOGNVFSCSVHREALHNYHTQKSLSPGK----- 228

Db 448 DGSFFLYSKLTVDKSRWQOGNVFSCSVHREALHNYHTQKSLSPGKVRSSRTSPDKP 507

QY 229 -----DWLK-----AFYDKVAEK----- 241

Db 508 VAHVANPQAEGLQWLNRRNALLANGVELNQLVVPSEGLYLIYSQVLFKGGCPST 567

QY 242 -----LK-----EAFDWLKAFY-----DKV 256

Db 568 HVLLTHITISIAVSYQTKVNLSSAIKSPCQRETPEGAEPWYEPYILGGVWFQLEKGDRL 627

QY 257 AEKL 260

Db 628 SAEI 631

RESULT 7

AAB28693

ID AAB28693 standard; Protein; 423 AA.

XX

AC AAB28693;

XX

DT 14-FEB-2001 (first entry)

XX

DE Fc-huAGP-1 (114-281) fusion protein.

XX

XX human; AGP-1; type II transmembrane protein; cytostatic; antiviral;  
KW antiinflammatory; hepatotropic; antiarteriosclerotic; anti-HIV; HIV;  
KW human immunodeficiency virus; apoptosis; proliferative disorder;  
KW cancer; hepatitis; acquired immunodeficiency syndrome; AIDS;  
KW autoimmune disorder; transplant rejection; cardiovascular disease;  
KW arteriosclerosis; Fc-huAGP-1; fusion protein.

XX

OS Homo sapiens.

XX

PN WO200063253-A1.

XX

PD 26-OCT-2000.

XX

PF 24-MAR-2000; 2000WO-US08004.

XX

PR 16-APR-1999; 99US-0293245.

XX

PA (AMGE-) AMGEN INC.

XX

PI Hsu H, Meng S;

XX

XX WPI: 2000-665240/64.

DR N-PSDB; AAC67833.

XX

PT Fusion protein of AGP-1 protein and an Fc region, used to treat  
PT proliferative disorders, immune disorders, and virally-induced  
PT disorders

XX

PS Disclosure; Fig 4; 93pp; English.

XX

XX The present sequence is an AGP-1 fusion protein. AGP-1 is a  
CC type II transmembrane protein. The fusion proteins comprise an Fc  
CC immunoglobulin region fused to the N-terminal portion of the AGP-1  
CC protein. The fusion proteins can be used to induce apoptosis in a tissue,  
CC and to treat proliferative disorders, immune disorders, or  
CC virally-induced disorders. The proliferative disorders include cancers,  
CC such as breast, prostate, lung or colon cancer. The viral infections  
CC include hepatitis, and acquired immunodeficiency syndrome (AIDS), and the  
CC immune disorders may be autoimmune disorders or transplant rejection.  
CC Cardiovascular diseases such as arteriosclerosis may also be treated. The  
CC AGP-1 containing fusion proteins have increased biological activity  
CC compared to the soluble AGP-1 proteins used in prior art therapies.

XX

SQ Sequence 423 AA;

Query Match 87.4%; Score 1250; DB 21; Length 423;  
Best Local Similarity 65.6%; Pred. No. 2.2e-33;  
Matches 240; Conservative 6; Mismatches 15; Indels 105; Gaps 5;

QY 2 DKHTCCPCPAPELLGGPSVFLPFPKPKDTLMISRTPEVTCVVVDVSHEDPEVKFNMYVD 61

Db 29 DKHTCCPCPAPELLGGPSVFLPFPKPKDTLMISRTPEVTCVVVDVSHEDPEVKFNMYVD 88

QY 62 GVEVHNAKTPREEQYNSTYRVVSVLTVLHQDWLNGKEYKCKVSNKALPAPIETISKAK 121

Db 89 GVEVHNAKTPREEQYNSTYRVVSVLTVLHQDWLNGKEYKCKVSNKALPAPIETISKAK 148

QY 122 GQPREPOVYLLPPSRDELTKNQVSLTCLVKGFYPSDIAVEWESNGQPENNYKTTTPPVLD 181

Db 149 GQPREPOVYLLPPSRDELTKNQVSLTCLVKGFYPSDIAVEWESNGQPENNYKTTTPPVLD 208

QY 182 DGSFFLYSKLTVDKSRWQOGNVFSCSVHREALHNYHTQKSLSPGK----- 228

```

Db 209 DGSFFLYSKLTVDKSRWQGNVFCSCVMHEALHNHYTKQSLSLSPGKVRERGQPVAAHI 268
QY 229 -----DW-----LKAFFY----- 235
Db 269 TGTRGRSNTLSSPNSKNEKALGRKINSWESSRSGHSFSLNLHURNGELVIHEKGFYIYS 328
QY 236 -----DKVAEKLKEAFDWLKAFY-----DKVAE----- 258
Db 329 QTVFRQEEIKENTKNDKQMVQIYKVTSPDPILLMKMSARNSCWSKDAEYGLYSIQGG 388
QY 259 --KLKE 262
Db 389 IFELKE 394

RESULT 8
AAB28692
ID AAB28692 standard; Protein; 441 AA.
XX
AC AAB28692;
DT 14-FEB-2001 (first entry)
DE Fc-huAGP-1 (95-281) fusion protein.
XX
KW Human; AGP-1; type II transmembrane protein; cytostatic; antiviral;
KW antiinflammatory; hepatotropic; antiarteriosclerotic; anti-HIV; HIV;
KW human immunodeficiency virus; apoptosis; proliferative disorder;
KW cancer; hepatitis; acquired immunodeficiency syndrome; AIDS;
KW autoimmune disorder; transplant rejection; cardiovascular disease;
KW arteriosclerosis; Fc-huAGP-1; fusion protein.
XX
OS Homo sapiens.
XX
PN WO200063253-A1.
XX
PD 26-OCT-2000.
XX
PF 24-MAR-2000; 2000WO-US08004.
XX
PR 16-APR-1999; 99US-0293245.
XX
PA (AMGE-) AMGEN INC.
XX
PI Hsu H, Meng S;
XX
DR WPI; 2000-665240/64.
DR N-PSDB; AAC67832.
XX
PT Fusion protein of AGP-1 protein and an Fc region, used to treat
PT proliferative disorders, immune disorders, and virally-induced
PT disorders -
XX
PS Disclosure; Fig 3; 93pp; English.
XX
CC The present sequence is an AGP-1 fusion protein. AGP-1 is a
CC type II transmembrane protein. The fusion proteins comprise an Fc
CC immunoglobulin region fused to the N-terminal portion of the AGP-1
CC protein. The fusion proteins can be used to induce apoptosis in a tissue,
CC and to treat proliferative disorders, immune disorders, or
CC virally-induced disorders. The proliferative disorders include cancers,
CC such as breast, prostate, lung or colon cancer. The viral infections
CC include hepatitis, and acquired immunodeficiency syndrome (AIDS), and the
CC immune disorders may be autoimmune disorders or transplant rejection.
CC Cardiovascular diseases such as arteriosclerosis may also be treated. The
CC AGP-1 containing fusion proteins have increased biological activity
CC compared to the soluble AGP-1 proteins used in prior art therapies.
XX
SQ Sequence 441 AA;
Query Match 87.4%; Score 1250; DB 21; Length 441;
Best Local Similarity 62.3%; Pred. No. 2.4e-33;
```

```

Matches 240; Conservative 6; Mismatches 15; Indels 124; Gaps 5;
QY 2 DKTHTCPPCPAPPELLGGPSVFLFPPPKPDTLMTSRTPVTCVVVDVSHEDPEVKFNWYVD 61
Db 29 DKTHTCPPCPAPPELLGGPSVFLFPPPKPDTLMTSRTPVTCVVVDVSHEDPEVKFNWYVD 88
QY 62 GVEVHNAKTKPREEQYNSTYRVVSVLTVLHQDLNGLNGKEYCKKVKSNKALPAPIEKTISKAK 121
Db 89 GVEVHNAKTKPREEQYNSTYRVVSVLTVLHQDLNGLNGKEYCKKVKSNKALPAPIEKTISKAK 148
QY 122 GQPREPOVYTLPPSRDELTKNQVSLTCLVKGFYPSDIAVEWESNGOPENNYKTTTPVLD 181
Db 149 GQPREPOVYTLPPSRDELTKNQVSLTCLVKGFYPSDIAVEWESNGOPENNYKTTTPVLD 208
QY 182 DGSFFLYSKLTVDKSRWQGNVFCSCVMHEALHNHYTKQSLSLSPGKVRERGQPVAAHI 228
Db 209 DGSFFLYSKLTVDKSRWQGNVFCSCVMHEALHNHYTKQSLSLSPGKVRERGQPVAAHI 268
QY 229 -----DW----- 230
Db 269 QNISPLVREGRGQPVAAHITGTRGRSNTLSSPNSKNEKALGRKINSWESSRSGHSFSLNL 328
QY 231 -----LKAFFY-----DKVAEKLKEAFDWLKAFY----- 253
Db 329 HLRNGELVIHEKGFYIYSQTVFRQEEIKENTKNDKQMVQIYKVTSPDPILLMKMSAR 388
QY 254 ----DKVAE-----KLKE 262
Db 389 NSCWSKDAEYGLYSIQGGIFELKE 413

RESULT 9
AAM52157
ID AAM52157 standard; Protein; 730 AA.
XX
AC AAM52157;
DT 05-FEB-2002 (first entry)
DE Humanised HMEF-1 heavy chain/DNase I fusion protein 2.
XX
KW Humanised monoclonal antibody; polymorphic epithelial mucin; PEM1;
KW cytotoxic; endonuclease; DNase I; human; cytostatic; cancer; apoptosis.
XX
OS Homo sapiens.
OS Synthetic.
XX
PN WO200174905-A1.
XX
PD 11-OCT-2001.
XX
PF 26-MAR-2001; 2001WO-GB01324.
XX
PR 03-APR-2000; 2000GB-0008049.
PR 02-OCT-2000; 2000US-237159P.
XX
PA (ANTI-) ANTISOMA RES LTD.
XX
PI Young RJ;
XX
WPI; 2001-662969/76.
XX
Novel compound used to treat cancer has target cell-specific portion
comprising humanised monoclonal antibody having specificity for
polymorphic epithelial mucin, and cytotoxic portion having
endonucleolytic activity -
XX
PS Claim 20; Figure 8; 176pp; English.
XX
CC The invention relates to a compound which comprises a target
CC cell-specific portion, comprising an humanised monoclonal antibody,
CC having specificity for polymorphic epithelial mucin (PEM) or its antigen
CC binding fragment and a cytotoxic portion having endonucleolytic activity,
```



CC exemplified by AAM52154-AAM52168 and encoded by ABA02682-ABA02728. The  
 CC compound has cytostatic activity useful for treating cancer and acting as  
 CC a potential inducer of apoptosis.

XX Sequence 730 AA;

Query Match 87.4%; Score 1250; DB 22; Length 730;  
 Best Local Similarity 50.4%; Pred. No. 7.8e-33;  
 Matches 241; Conservative 6; Mismatches 15; Indels 216; Gaps 6;

QY 2 DKTHTCPPCPAPPELLGGPSVFLFPKPKDTLMISRTPEVTCVVDVSHEDPEVKFNWYVD 61  
 |||||  
 DB 241 DKTHTCPPCPAPPELLGGPSVFLFPKPKDTLMISRTPEVTCVVDVSHEDPEVKFNWYVD 300  
 |||||  
 QY 62 GVEVHNAKTPREEQNSTYRVVSVLTVLHODWLNKGYCKVSNKALPAPIEKTISKAK 121  
 |||||  
 DB 301 GVEVHNAKTPREEQNSTYRVVSVLTVLHODWLNKGYCKVSNKALPAPIEKTISKAK 360  
 |||||  
 QY 122 GQPREQVYTLPPSRDELTKNQVSLTCLVKGFYPSDIAVEWESNGQPENNYKTTTPPVLD 181  
 |||||  
 DB 361 GQPREQVYTLPPSRDELTKNQVSLTCLVKGFYPSDIAVEWESNGQPENNYKTTTPPVLD 420  
 |||||  
 QY 182 DGSFFLYSKLTVDKSRWQGNVFSCVMHEALHNHYTQKSLSLSP-GKDWLK----- 232  
 |||||  
 DB 421 DGSFFLYSKLTVDKSRWQGNVFSCVMHEALHNHYTQKSLSLSPGKGLKIAAFNIQT 480  
 |||||  
 QY 233 -----KAFYDKV-----AEKLKEA 263  
 |||||  
 DB 661 TFQWLIPDSADTTATPTHCAIDRVVAGMLLRGAVVPDSALPFFNQAYGLSDQLAQ 718  
 |||||

RESULT 10  
 AAM52160  
 ID AAM52160 standard; Protein; 740 AA.  
 XX  
 AC AAM52160;  
 XX  
 DT 05-FEB-2002 (first entry)  
 XX  
 DE Humanised HMFG-1 heavy chain/DNase I fusion protein 5.  
 XX  
 DE Humanised monoclonal antibody; polymorphic epithelial mucin; PEM1;  
 KW cytotoxic; endonuclease; DNase I; human; cytostatic; cancer; apoptosis.  
 XX  
 OS Homo sapiens.  
 OS Synthetic.  
 XX  
 PN WO200174905-A1.  
 XX  
 PD 11-OCT-2001.  
 XX  
 PF 26-MAR-2001; 2001WO-GB01324.  
 XX  
 PR 03-APR-2000; 2000GB-0008049.  
 PR 02-OCT-2000; 2000US-237159P.  
 XX  
 PA (ANTI-) ANTISOMA RES LTD.  
 XX  
 PI Young RJ;  
 XX  
 DR WPI; 2001-662969/76.

XX Novel compound used to treat cancer has target cell-specific portion  
 PT comprising humanised monoclonal antibody having specificity for  
 PT polymorphic epithelial mucin, and cytotoxic portion having  
 PT endonucleolytic activity

XX Claim 20; Figure 11; 176pp; English.

XX The invention relates to a compound which comprises a target  
 CC cell-specific portion, comprising a humanised monoclonal antibody,  
 CC having specificity for polymorphic epithelial mucin (PEM) or its antigen  
 CC binding fragment and a cytotoxic portion having endonucleolytic activity,  
 CC exemplified by AAM52154-AAM52168 and encoded by ABA02682-ABA02728. The  
 CC compound has cytostatic activity useful for treating cancer and acting as  
 CC a potential inducer of apoptosis.

XX Sequence 740 AA;

Query Match 87.4%; Score 1250; DB 22; Length 740;  
 Best Local Similarity 50.4%; Pred. No. 8e-33;  
 Matches 241; Conservative 6; Mismatches 15; Indels 216; Gaps 6;

QY 2 DKTHTCPPCPAPPELLGGPSVFLFPKPKDTLMISRTPEVTCVVDVSHEDPEVKFNWYVD 61  
 |||||  
 DB 241 DKTHTCPPCPAPPELLGGPSVFLFPKPKDTLMISRTPEVTCVVDVSHEDPEVKFNWYVD 300  
 |||||  
 QY 62 GVEVHNAKTPREEQNSTYRVVSVLTVLHODWLNKGYCKVSNKALPAPIEKTISKAK 121  
 |||||  
 DB 301 GVEVHNAKTPREEQNSTYRVVSVLTVLHODWLNKGYCKVSNKALPAPIEKTISKAK 360  
 |||||  
 QY 122 GQPREQVYTLPPSRDELTKNQVSLTCLVKGFYPSDIAVEWESNGQPENNYKTTTPPVLD 181  
 |||||  
 DB 361 GQPREQVYTLPPSRDELTKNQVSLTCLVKGFYPSDIAVEWESNGQPENNYKTTTPPVLD 420  
 |||||  
 QY 182 DGSFFLYSKLTVDKSRWQGNVFSCVMHEALHNHYTQKSLSLSP-GKDWLK----- 232  
 |||||  
 DB 421 DGSFFLYSKLTVDKSRWQGNVFSCVMHEALHNHYTQKSLSLSPGKGLKIAAFNIQT 480  
 |||||  
 QY 233 -----KAFYDKV-----AEKLKEA 263  
 |||||  
 DB 661 TFQWLIPDSADTTATPTHCAIDRVVAGMLLRGAVVPDSALPFFNQAYGLSDQLAQ 718  
 |||||

RESULT 11  
 AAY84965  
 ID AAY84965 standard; Protein; 633 AA.  
 XX  
 AC AAY84965;  
 XX  
 DT 21-AUG-2000 (first entry)  
 XX  
 DE Amino acid sequence of a CD-20 specific chimeric receptor.

XX CD20-specific receptor; CD-20 specific redirected T cell; leukemia;  
 KW CD20+ malignancy; non-Hodgkin's lymphoma; myeloablative chemotherapy;  
 KW stem cell rescue; autoimmune disease; lupus; rheumatoid arthritis.  
 XX  
 OS Synthetic.  
 OS Mus sp.  
 XX  
 FX Key Location/Qualifiers

FT	Peptide	27..86
FT	/note= "signal peptide from murine T86.66 antibody	
FT	kappa light chain"	
FT	Region	21..126
FT	/note= "kappa variable regions"	
FT	Peptide	145..266
FT	/note= "GS18 linker"	
FT	Region	283..392
FT	/note= "hinge region"	
FT	Region	393..499
FT	/note= "CH3 region"	
FT	Region	500..521
FT	/note= "CD4 transmembrane region"	
FT	Region	522..633
FT	/note= "zeta chain"	
XX	WO200023573-A2.	
XX	PN	
XX	PD	27-APR-2000.
XX	PF	20-OCT-1999; 99WO-US24484.
XX	PR	20-OCT-1998; 98US-0105014.
XX	(CITY ) CITY OF HOPE.	
PA	Raubitschek A, Jensen MC, Wu AM;	
PI	WPI; 2000-339676/29.	
DR	N-PSDB; AAA15019.	
XX	Genetically engineered CD20-specific redirected T cells useful for	
PT	treating a CD20+ malignancy, such as non-Hodgkin's lymphoma or CD20+	
PT	acute or chronic leukemia, and autoimmune disease -	
XX	Claim 10; Page 53-55; 58pp; English.	
XX	The present sequence represents a synthetic CD20-specific chimeric	
CC	receptor. The specification describes CD-20 specific redirected T cells	
CC	which express and bear on the cell surface membrane a CD20-chimeric	
CC	receptor comprising an intracellular signalling domain, a transmembrane	
CC	domain and an extracellular domain, the extracellular domain comprising	
CC	a CD20-specific receptor. The genetically engineered CD20-specific	
CC	redirected T cells are useful for treating a CD20+ malignancy, such	
CC	as non-Hodgkin's lymphoma or CD20+ acute or chronic leukemia, in a	
CC	human patient having previously undergone myeloablative chemotherapy and	
CC	stem cell rescue. The genetically engineered CD20-specific redirected	
CC	T cells are also useful for abrogating an untoward B cell function, such	
CC	as autoimmune disease (lupus or rheumatoid arthritis) in a patient.	
XX	Sequence 633 AA;	
SQ	Query Match 87.3%; Score 1249; DB 21; Length 633;	
	Best Local Similarity 73.9%; Pred. No. 6.1e-33;	
	Matches 238; Conservative 7; Mismatches 16; Indels 61; Gaps	
QY	2 DKHTCPPCPAPELLGSPVFLPPPKD <sup>1</sup> LTMSIRPTEVTVCVVVDVSHEDPEVKFNWYVD 61	
Db		
	273 DKHTCPPCPAPELLGSPVFLPPPKD <sup>1</sup> LTMSIRPTEVTVCVVVDVSHEDPEVKFNWYVD 332	
QY	62 GVEVHNAKTPREQ <sup>1</sup> NSYTRVSVLTVLHQDWLNGKEYCKCKVNKALPAPIEKTISKAK 121	
Db		
	333 GVEVHNAKTPREQ <sup>1</sup> NSYTRVSVLTVLHQDWLNGKEYCKCKVNKALPAPIEKTISKAK 392	
QY	122 GQPREPOVYTLPPSRDELTKNQVSLTCLVKGFYPSDIAVEWESNGOPENNYKTTTPPVLD 181	
Db		
	393 GQPREPOVYTLPPSRDELTKNQVSLTCLVKGFYPSDIAVEWESNGOPENNYKTTTPPVLD 452	
QY	182 DGSFELYSKLTVDKSRWQQGNFSCSMHEALHNNHYTQTLSLSFGKDWL----- 231	
Db		
	453 DGSFFLYSKLTVDKSRWQQGNFSCSMHEALHNNHYTQTLSLSFGKDWL----- 512	
QY	232 -----KATYDKVAEKLEKAFAWL----- 249	

Db 89 GVEVHNAKTPREYOYNSTYRVVSVLTVLHQDWLNGKEYKCKVSNKALPAPIEKTISKAK 148  
QY 122 GQPREPOVYTLPPSRDELTKNQVSLTCLVKGFYPSDIAVEWESNGQPENNYKTTTPPVLD 181  
Db 149 GQPREPOVYTLPPSRDELTKNQVSLTCLVKGFYPSDIAVEWESNGQPENNYKTTTPPVLD 208  
QY 182 DGSFFLYSKLTVDKSRWQOGNVSFSCSYMHEALHNHYTQKSLSLSPGK----- 228  
Db 209 DGSFFLYSKLTVDKSRWQOGNVSFSCSYMHEALHNHYTQKSLSLSPGKGRPKVAHITG 268  
QY 229 -----DW-----LKAIFYDKV 238  
Db 269 ITRRSNALIPISKDGTLCQKIESWESSRKGHSFLNHVLFNRGELVIEQEGLYIYSQT 328  
QY 239 AEKLKEAFDWLK-AFYDKVAEK-----LKEA 263  
Db 329 YFRFOEAEDASKMVKDKVTKQLVQYIYKTSYTPDPVILMKSA 372  
RESULT 13  
AAB28694  
ID AAB28694 standard; Protein; 448 AA.  
XX  
AC AAB28694;  
XX  
DT 14-FEB-2001 (first entry)  
DE FC-muAGP-1 (99-291) fusion protein.  
XX  
KW Mouse; AGP-1; type II transmembrane protein; cytostatic; antiviral;  
KW antiinflammatory; hepatotropic; antiarteriosclerotic; anti-HIV; HIV;  
KW human immunodeficiency virus; apoptosis; proliferative disorder;  
KW cancer; hepatitis; acquired immunodeficiency syndrome; AIDS;  
KW autoimmune disorder; transplant rejection; cardiovascular disease;  
KW arteriosclerosis; FC-muAGP-1; fusion protein.  
XX  
OS Mus sp.  
XX  
PN WO200063253-A1.  
XX  
PD 26-OCT-2000.  
XX  
PF 24-MAR-2000; 2000WO-US08004.  
XX  
PR 16-APR-1999; 99US-0293245.  
XX  
PA (AMGE-) AMGEN INC.  
XX  
PI Hsu H, Meng S;  
XX  
DR WPI; 2000-665240/64.  
DR N-PSDB; AAC67834.  
XX  
PT Fusion protein of AGP-1 protein and an Fc region, used to treat  
PT proliferative disorders, immune disorders, and virally-induced  
PT disorders -  
XX  
PS Disclosure; Fig 5; 93pp; English.  
XX  
CC The present sequence is an AGP-1 fusion protein. AGP-1 is a  
CC type II transmembrane protein. The fusion proteins comprise an Fc  
CC immunoglobulin region fused to the N-terminal portion of the AGP-1  
CC protein. The fusion proteins can be used to induce apoptosis in a tissue,  
CC and to treat proliferative disorders, immune disorders, or  
CC virally-induced disorders. The proliferative disorders include cancers,  
CC such as breast, prostate, lung or colon cancer. The viral infections  
CC include hepatitis, and acquired immunodeficiency syndrome (AIDS), and the  
CC immune disorders may be autoimmune disorders or transplant rejection.  
CC Cardiovascular diseases such as arteriosclerosis may also be treated. The  
CC AGP-1 containing fusion proteins have increased biological activity  
CC compared to the soluble AGP-1 proteins used in prior art therapies.  
XX  
SQ Sequence 448 AA;

Query Match 87.2%; Score 1247; DB 21; Length 448;  
Best Local Similarity 65.8%; Pred. No. 3.2e-33;  
Matches 240; Conservative 4; Mismatches 18; Indels 103; Gaps 4;  
QY 2 DKTHTCPPCPAPELGGPSVFLPPPKDPTLMISRTPEVTCVVVDVSHEDPEVKFNKYVD 61  
Db 29 DKTHTCPPCPAPELGGPSVFLPPPKDPTLMISRTPEVTCVVVDVSHEDPEVKFNKYVD 88  
QY 62 GVEVHNAKTPREYOYNSTYRVVSVLTVLHQDWLNGKEYKCKVSNKALPAPIEKTISKAK 121  
Db 89 GVEVHNAKTPREYOYNSTYRVVSVLTVLHQDWLNGKEYKCKVSNKALPAPIEKTISKAK 148  
QY 122 GQPREPOVYTLPPSRDELTKNQVSLTCLVKGFYPSDIAVEWESNGQPENNYKTTTPPVLD 181  
Db 149 GQPREPOVYTLPPSRDELTKNQVSLTCLVKGFYPSDIAVEWESNGQPENNYKTTTPPVLD 208  
QY 182 DGSFFLYSKLTVDKSRWQOGNVSFSCSYMHEALHNHYTQKSLSLSPGK----- 228  
Db 209 DGSFFLYSKLTVDKSRWQOGNVSFSCSYMHEALHNHYTQKSLSLSPGKGRPKVAHITG 268  
QY 229 -----DW-----LKAIFYDKV 238  
Db 269 LSTPPLPRGGRPKVAHITGITRRSNSALIPISKDGTLCQKIESWESSRKGHSFLNHV 328  
QY 231 -----LKAIFYDKVAEKLEAFDWLK-AFYDKVAEK----- 259  
Db 329 LFRNGELVIEQEGLYIYSQTYFRFOEAEDASKMVKDKVTKQLVQYIYKTSYTPDPV 388  
QY 260 -LKEA 263  
Db 389 LMKSA 393  
RESULT 14  
AAB16958  
ID AAB16958 standard; Protein; 247 AA.  
XX  
AC AAB16958;  
XX  
DT 31-OCT-2000 (first entry)  
DE FC-TMP protein sequence SEQ ID NO:6.  
XX  
KW Modified peptide; therapeutic agent; fusion; Fc domain; cancer;  
KW autoimmune disease; cytostatic; antiasthmatic; thrombolytic; VEGF;  
KW immunosuppressive; EPO; TPO; CTLA4; mimetic; IL-1; TNF; antagonist;  
KW MMP; inhibitor; erythropoietin; thrombopoietin; interleukin 1;  
KW cytotoxic T cell lymphocyte antigen 4; tumour necrosis factor;  
KW vascular endothelial growth factor; matrix metalloproteinase;  
KW asthma; thrombosis; pharmaceutical.  
XX  
OS Homo sapiens.  
OS Synthetic.  
XX  
PN WO200024782-A2.  
XX  
PD 04-MAY-2000.  
XX  
PF 25-OCT-1999; 99WO-US25044.  
XX  
PR 23-OCT-1998; 98US-0105371.  
PR 22-OCT-1999; 99US-0428082.  
XX  
PA (AMGE-) AMGEN INC.  
XX  
PI Feige U, Liu C, Cheetham J, Boone TC;  
XX WPI; 2000-350702/30.  
DR N-PSDB; AAA69444.  
XX  
PT Novel composition of matter comprising an Fc domain and  
PT pharmacologically active peptides, useful for treating cancer and

Claim 21; Page 179-180; 608pp; English.

Sequence	247 AA:
SQ	

RESULT 15

AC ABB73411:

Fc-TPO mimetic peptide (Fc-TMP) amino acid SEQ ID NO:6.

Modified peptide; mimetic; Fc domain; fusion; immunoglobulin G; IgG; EPO; erythropoietin; TPO; tumour necrosis factor alpha inhibitor; TNF-alpha inhibitor; interleukin 1 antagonist; IL-1 antagonist; TMP; TPO mimetic peptide; EPO mimetic peptide; EMP; VEGF antagonist; MMP inhibitor; antiinflammatory; antitumour; immunosuppressive; cytostatic; antirheumatic; antiarthritis; antidiabetic; ophthalmological; antianemic; anorectic; antiinfectivity; autoimmune disease; dermatological; neuroprotective; inflammatory disease; haemostatic disease; tumour growth; cancer; rheumatoid arthritis; diabetic retinopathy; infertility; obesity; sleep disorder; neurological degenerative disease; anaemia; thrombocytopaenia; metastatic tumour; systemic lupus erythematosus; Fanconi's syndrome.

XX  
XX

Search completed: April 21, 2003, 10:45:28  
Job time.: 33.0588 secs

---



GenCore version 5.1.4\_p5\_4578  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: April 21, 2003, 10:42:26 ; Search time 14.7529 Seconds  
(without alignments)  
526.515 Million cell updates/sec

Title: 2LINK7LINK7  
Perfect score: 1430  
Sequence: 1 MDKHTCPCPCAPPELLGGPS.....AFDWLKAFYDKVAEKLKEAF 264

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.0

Searched: 262574 seqs, 29422922 residues

Total number of hits satisfying chosen parameters: 262574

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Issued\_Patents\_AA.\*  
1: /cgn2.6/ptodata/1/iaa/5A\_COMB.pep.\*  
2: /cgn2.6/ptodata/1/iaa/5B\_COMB.pep.\*  
3: /cgn2.6/ptodata/1/iaa/6A\_COMB.pep.\*  
4: /cgn2.6/ptodata/1/iaa/6B\_COMB.pep.\*  
5: /cgn2.6/ptodata/1/iaa/PCUS\_COMB.pep.\*  
6: /cgn2.6/ptodata/1/iaa/backfiles.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	ID	Description
1	1240	86.7	389	US-09-131-247-14	Sequence 14, Appl
2	1239	86.6	660	US-09-181-706-8	Sequence 8, Appl
3	1239	86.6	660	US-09-458-791-8	Sequence 8, Appl
4	1239	86.6	660	US-09-459-066-8	Sequence 8, Appl
5	1234	86.3	347	US-07-940-861-43	Sequence 43, Appl
6	1234	86.3	347	US-08-459-512-43	Sequence 43, Appl
7	1234	86.3	347	US-08-459-657-43	Sequence 43, Appl
8	1234	86.3	347	US-08-460-132-43	Sequence 43, Appl
9	1234	86.3	347	US-08-466-465-8	Sequence 8, Appl
10	1234	86.3	347	PCT-US92-02050-43	Sequence 43, Appl
11	1234	86.3	482	US-09-189-129-2	Sequence 2, Appl
12	1233	86.2	232	US-08-595-043A-50	Sequence 50, Appl
13	1233	86.2	235	US-09-131-247-6	Sequence 6, Appl
14	1233	86.2	331	US-09-178-869-2	Sequence 2, Appl
15	1233	86.2	360	US-09-180-100-11	Sequence 11, Appl
16	1233	86.2	371	US-08-236-311-7	Sequence 7, Appl
17	1233	86.2	371	US-08-457-918-7	Sequence 7, Appl
18	1233	86.2	376	US-09-180-100-22	Sequence 22, Appl
19	1233	86.2	387	US-08-470-299-4	Sequence 4, Appl
20	1233	86.2	388	US-09-131-247-16	Sequence 16, Appl
21	1233	86.2	396	US-08-784-512-3	Sequence 3, Appl
22	1233	86.2	396	US-09-176-228-3	Sequence 3, Appl
23	1233	86.2	424	US-09-333-593A-8	Sequence 8, Appl
24	1233	86.2	424	PCT-US95-03866-12	Sequence 12, Appl
25	1233	86.2	424	PCT-US95-03866-14	Sequence 14, Appl
26	1233	86.2	437	PCT-US96-10043-11	Sequence 11, Appl
27	1233	86.2	442	PCT-US96-10043-9	Sequence 9, Appl

28	1233	86.2	446	3	US-08-397-411-7	Sequence 7, Appl
29	1233	86.2	449	1	US-08-458-516-13	Sequence 13, Appl
30	1233	86.2	459	1	US-08-157-101A-7	Sequence 7, Appl
31	1233	86.2	476	2	US-08-378-939-10	Sequence 10, Appl
32	1233	86.2	476	3	US-08-487-550-4	Sequence 4, Appl
33	1233	86.2	476	3	US-08-487-550-12	Sequence 12, Appl
34	1233	86.2	478	3	US-08-487-550-8	Sequence 8, Appl
35	1233	86.2	488	4	US-08-776-511-2	Sequence 2, Appl
36	1233	86.2	680	4	US-08-227-496C-15	Sequence 15, Appl
37	1233	86.2	711	4	US-09-485-737B-90	Sequence 90, Appl
38	1231	86.1	254	2	US-08-284-391B-33	Sequence 33, Appl
39	1231	86.1	254	4	US-09-218-950-33	Sequence 33, Appl
40	1228	85.9	664	3	US-08-957-063-16	Sequence 16, Appl
41	1228	85.9	664	3	US-09-957-063-18	Sequence 18, Appl
42	1228	85.9	664	4	US-09-487-685-16	Sequence 16, Appl
43	1228	85.9	664	4	US-09-487-685-18	Sequence 18, Appl
44	1228	85.9	664	4	US-08-802-805D-16	Sequence 16, Appl
45	1228	85.9	664	4	US-08-802-805D-18	Sequence 18, Appl

ALIGNMENTS

RESULT 1  
US-09-131-247-14  
; Sequence 14, Application US/09131247  
; Patent No. 6294170  
; GENERAL INFORMATION:  
; APPLICANT: Boone, Thomas C.  
; APPLICANT: Herhenson, Susan  
; APPLICANT: Bevilacqua, Michael P.  
; APPLICANT: Collins, David S.  
; TITLE OF INVENTION: COMPOSITION AND METHOD FOR TREATING INFLAMMATORY  
; FILE REFERENCE: A-365F  
; CURRENT APPLICATION NUMBER: US/09/131,247  
; CURRENT FILING DATE: 1998-08-07  
; EARLIER APPLICATION NUMBER: 60/055,185  
; EARLIER FILING DATE: 1997-08-08  
; EARLIER APPLICATION NUMBER: PCT/US 97/02131  
; EARLIER FILING DATE: 1997-02-10  
; NUMBER OF SEQ ID NOS: 16  
; SOFTWARE: Patent In Ver. 2.0  
; SEQ ID NO 14  
; LENGTH: 389  
; TYPE: PRT  
; ORGANISM: Human  
US-09-131-247-14

Query Match	86.7%	Score 1240;	DB 4;	Length 389;
Best Local Similarity	63.7%	Pred. No. 8.4e-40;		
Matches	239;	Conservative	5;	Mismatches 8;
				Indels 123;
				Gaps 7;
QY	2	DKHTCPCPCAPPELLGGPSVFLFPKPKDTLMISRTPEVTCVVVDVSHEDPEVKFNWYVD	61	
Db	10	DKHTCPCPCAPPELLGGPSVFLFPKPKDTLMISRTPEVTCVVVDVSHEDPEVKFNWYVD	69	
QY	62	GVEVHNAKTPREOVNSTRYVSVLTVLHODWLNKGYCKVSNKALPAPIETISKAK	121	
Db	70	GVEVHNAKTPREOVNSTRYVSVLTVLHODWLNKGYCKVSNKALPAPIETISKAK	129	
QY	122	GQPREQVYTLPPSRDELTKNQVSLTCLVKGFYPSDIAVEWESNGQPENNYKTPPVLD	181	
Db	130	GQPREQVYTLPPSRDELTKNQVSLTCLVKGFYPSDIAVEWESNGQPENNYKTPPVLD	189	
QY	182	DGSFFLYSKLTVDKSRWQQGNVFCSCVMHEALNHHYTKQSLSLSPGK	228	
Db	190	DGSFFLYSKLTVDKSRWQQGNVFCSCVMHEALNHHYTKQSLSLSPGKMRPSGRKSKMQA	249	
QY	229	--DW---LKAFY-----DKV-----	238	
Db	250	FRWVDNQKTFYLRNNOLVAGYLGQPNVLEEKIDVVPPIEPHALFIHGGKMLSCVKS	309	

QY	2	DKTHTCPPCPAPDELGGPSVFLFPKPKDPTLMISRTPEVTCVVVDVSHPEDDEVKNKYVD	61
Dd	37	DKTHTCPPCPAPAEAGASVFLFPKPKDPTLMISRTPEVTCVVVDVSHPEDDEVKNKYVD	96
QY	62	GVEVHNNAKTPREQYNSTRYVSVLTVLHQDLNGKEYCKVSNKALPAPIETKISKAK	121
Dd	97	GVEVHNNAKTPREQYNSTRYVSVLTVLHQDLNGKEYCKVSNKALPAPIETKISKAK	156
QY	122	GGPREQVYTLPPSRDELTKNQVSLTCLVKGPYPSSDIVAESWESNGQPENNYKTTTPPVLDLS	181

Query Match	86.68;	Score 1239;	DB 4;	Length 660;
-------------	--------	-------------	-------	-------------



```
Best Local Similarity 48.5%; Pred. No. 3.3e-39;
Matches 238; Conservative 7; Mismatches 18; Indels 228; Gaps 6;

QY 2 DKHTCTPCPAPELLGGPSVFLFPPKPKDTLMISRTPEVTCVVVDVSHEDPEVKFNWYVD 61
Db 37 DKHTCTPCPAPEAGPSVFLFPPKPKDTLMISRTPEVTCVVVDVSHEDPEVKFNWYVD 96
QY 62 GVEVHNAKTPREQYNSTYRVVSVLTVLHQDWLNGKEYKCKVSNKALPAPIEKTISKAK 121
Db 97 GVEVHNAKTPREQYNSTYRVVSVLTVLHQDWLNGKEYKCKVSNKALPAPIEKTISKAK 156
QY 122 GQPREPOVYTLPPSRDELTKNQVSLTCLVKGFYPSDIAVEWESNGQPENNYKTTTPPVLD 181
Db 157 GQPREPOVYTLPPSRDELTKNQVSLTCLVKGFYPSDIAVEWESNGQPENNYKTTTPPVLD 216
QY 182 DGSFFLYSKLTVDKSRQOGNVFSCSVMEALHNNHYTKQSLSPGK----- 228
Db 217 DGSFFLYSKLTVDKSRQOGNVFSCSVMEALHNNHYTKQSLSPGKGGGGGGGSGT 276
QY 229 DWLK----- 232
Db 277 EWHKFETSEELISTYLIDDLVLYTCVNGAVYTFSSNNELNKTGLTNNNNYITTSIKVEDTLV 336
QY 233 ----- 232
Db 337 CGTNGNPKCKWIDGSEDPKRYGRGYAPYQNSKVTIISHNECVLSINISKEGKRWRRF 396
QY 233 -----AF-----YDKVAEKLKEAFD----- 247
Db 397 DGPCGYDLYTADNVIPKDGVRGAFVDKGYDKVYILFTDTIDTKRIVKIPYIAQMC LND 456
QY 248 -----W-----LKAIFYDKVAEK----- 259
Db 457 EGGPSSLSHRSWTFELKVELECDIDGRSYROIHSKAIKTDNDTILVFFDSPYSKALC 516
QY 260 -----LKEAF 264
Db 517 TYSMNAIKHSF 527

RESULT 4
US-09-459-066-8
; Sequence 8, Application US/09459066
; Patent No. 6187909
; GENERAL INFORMATION:
; APPLICANT: Spriggs, Melanie
; TITLE OF INVENTION: VIRAL ENCODED SEMAPHORIN PROTEIN
; TITLE OF INVENTION: RECEPTOR DNA AND POLYPEPTIDES
; NUMBER OF SEQUENCES: 10
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Janis C. Henry
; STREET: 51 University St.
; CITY: Seattle
; STATE: WA
; COUNTRY: US
; ZIP: 98101
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: MS-DOS/Windows 95
; SOFTWARE: Word for Windows 95, 7.0a
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/459,066
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/958,598
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Henry, Janis C
; REGISTRATION NUMBER: 34,347
; REFERENCE/DOCKET NUMBER: 2631
; TELECOMMUNICATION INFORMATION:
```

```
TELEPHONE: (206)470-4189
TELEFAX: (206)233-0644
; INFORMATION FOR SEQ ID NO: 8:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 660 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-09-459-066-8

Query Match 86.6%; Score 1239; DB 4; Length 660;
Best Local Similarity 48.5%; Pred. No. 3.3e-39;
Matches 238; Conservative 7; Mismatches 18; Indels 228; Gaps 6;

QY 2 DKHTCTPCPAPELLGGPSVFLFPPKPKDTLMISRTPEVTCVVVDVSHEDPEVKFNWYVD 61
Db 37 DKHTCTPCPAPEAGPSVFLFPPKPKDTLMISRTPEVTCVVVDVSHEDPEVKFNWYVD 96
QY 62 GVEVHNAKTPREQYNSTYRVVSVLTVLHQDWLNGKEYKCKVSNKALPAPIEKTISKAK 121
Db 97 GVEVHNAKTPREQYNSTYRVVSVLTVLHQDWLNGKEYKCKVSNKALPAPIEKTISKAK 156
QY 122 GQPREPOVYTLPPSRDELTKNQVSLTCLVKGFYPSDIAVEWESNGQPENNYKTTTPPVLD 181
Db 157 GQPREPOVYTLPPSRDELTKNQVSLTCLVKGFYPSDIAVEWESNGQPENNYKTTTPPVLD 216
QY 182 DGSFFLYSKLTVDKSRQOGNVFSCSVMEALHNNHYTKQSLSPGK----- 228
Db 217 DGSFFLYSKLTVDKSRQOGNVFSCSVMEALHNNHYTKQSLSPGKGGGGGGGSGT 276
QY 229 DWLK----- 232
Db 277 EWHKFETSEELISTYLIDDLVLYTCVNGAVYTFSSNNELNKTGLTNNNNYITTSIKVEDTLV 336
QY 233 ----- 232
Db 337 CGTNGNPKCKWIDGSEDPKRYGRGYAPYQNSKVTIISHNECVLSINISKEGKRWRRF 396
QY 233 -----AF-----YDKVAEKLKEAFD----- 247
Db 397 DGPCGYDLYTADNVIPKDGVRGAFVDKGYDKVYILFTDTIDTKRIVKIPYIAQMC LND 456
QY 248 -----W-----LKAIFYDKVAEK----- 259
Db 457 EGGPSSLSHRSWTFELKVELECDIDGRSYROIHSKAIKTDNDTILVFFDSPYSKALC 516
QY 260 -----LKEAF 264
Db 517 TYSMNAIKHSF 527

RESULT 5
US-07-940-861-43
; Sequence 43, Application US/07940861
; Patent No. 5547853
; GENERAL INFORMATION:
; APPLICANT: BIOGEN, INC.
; APPLICANT: WALLNER, Barbara P.
; APPLICANT: MILLER, Glenn T.
; APPLICANT: ROSA, Margaret D.
; TITLE OF INVENTION: CD2-BINDING DOMAIN OF LYMPHOCYTE
; TITLE OF INVENTION: FUNCTION ASSOCIATED ANTIGEN 3
; NUMBER OF SEQUENCES: 43
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Fish & Neave
; STREET: 875 Third Avenue
; CITY: New York
; STATE: New York
; COUNTRY: U.S.A.
; ZIP: 10022-6250
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
```

OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/07/940,861  
FILING DATE: 21-OCT-1992  
CLASSIFICATION: 514  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: PCT/US92/02050  
FILING DATE: 12-MAR-1992  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/667,971  
FILING DATE: 12-MAR-1991  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/770,967  
FILING DATE: 07-OCT-1991  
ATTORNEY/AGENT INFORMATION:  
NAME: HALEY, James F., Jr.  
REGISTRATION NUMBER: 27,794  
REFERENCE/DOCKET NUMBER: B151CIP2  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (212)715-0600  
TELEFAX: (212)715-0673  
TELEX: 14-8367  
INFORMATION FOR SEQ ID NO: 43:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 347 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-07-940-861-43

Query Match 86.3%; Score 1234; DB 1; Length 347;  
Best Local Similarity 99.6%; Pred. No. 1.1e-39;  
Matches 227; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 MDKTHCPCPAPPELLGGSVLFPKPKDTLMISRTPEVTCVVDVSHEDPEVKFNWYV 60  
Db :|||||  
Db 120 VDKTHCPCPAPPELLGGSVLFPKPKDTLMISRTPEVTCVVDVSHEDPEVKFNWYV 179  
QY 61 DGVEVHNATKPREQYNSTYRVSVLTVLHODWLNKGEYCKVSNKALPAPIETKISK 120  
Db :|||||  
Db 180 DGVEVHNATKPREQYNSTYRVSVLTVLHODWLNKGEYCKVSNKALPAPIETKISK 239  
QY 121 KGQPREQYVTLPPSRDELTKNQVSLTCLVKGFYPSDIAVEWESNGQPENNYKTTTPPVLD 180  
Db :|||||  
Db 240 KGQPREQYVTLPPSRDELTKNQVSLTCLVKGFYPSDIAVEWESNGQPENNYKTTTPPVLD 299  
QY 181 SDGSFELYSLTVDKSRWQGNVFCSCVMHEALHNNHYTOKSLSPGK 228  
Db :|||||  
Db 300 SDGSFELYSLTVDKSRWQGNVFCSCVMHEALHNNHYTOKSLSPGK 347

RESULT 6  
US-08-459-512-43  
; Sequence 43, Application US/08459512  
; Patent No. 5728677  
; GENERAL INFORMATION:  
; APPLICANT: BIOGEN, INC.  
; APPLICANT: WALLNER, Barbara P.  
; APPLICANT: MILLER, Glenn T.  
; APPLICANT: ROSA, Margaret D.  
; TITLE OF INVENTION: CD2-BINDING DOMAIN OF LYMPHOCYTE  
; NUMBER OF SEQUENCES: 43  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Fish & Neave  
; CITY: New York  
; STATE: New York  
; COUNTRY: U.S.A.  
; ZIP: 10022-6250  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/459,512  
FILING DATE: 02-JUN-1995  
CLASSIFICATION: 514  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: PCT/US92/02050  
FILING DATE: 12-MAR-1992  
APPLICATION NUMBER: US 07/667,971  
FILING DATE: 12-MAR-1991  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/770,967  
FILING DATE: 07-OCT-1991  
ATTORNEY/AGENT INFORMATION:  
NAME: HALEY, James F., Jr.  
REGISTRATION NUMBER: 27,794  
REFERENCE/DOCKET NUMBER: B151CIP2  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (212)715-0600  
TELEFAX: (212)715-0673  
TELEX: 14-8367  
INFORMATION FOR SEQ ID NO: 43:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 347 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-459-512-43

Query Match 86.3%; Score 1234; DB 1; Length 347;  
Best Local Similarity 99.6%; Pred. No. 1.1e-39;  
Matches 227; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 MDKTHCPCPAPPELLGGSVLFPKPKDTLMISRTPEVTCVVDVSHEDPEVKFNWYV 60  
Db :|||||  
Db 120 VDKTHCPCPAPPELLGGSVLFPKPKDTLMISRTPEVTCVVDVSHEDPEVKFNWYV 179  
QY 61 DGVEVHNATKPREQYNSTYRVSVLTVLHODWLNKGEYCKVSNKALPAPIETKISK 120  
Db :|||||  
Db 180 DGVEVHNATKPREQYNSTYRVSVLTVLHODWLNKGEYCKVSNKALPAPIETKISK 239  
QY 121 KGQPREQYVTLPPSRDELTKNQVSLTCLVKGFYPSDIAVEWESNGQPENNYKTTTPPVLD 180  
Db :|||||  
Db 240 KGQPREQYVTLPPSRDELTKNQVSLTCLVKGFYPSDIAVEWESNGQPENNYKTTTPPVLD 299  
QY 181 SDGSFELYSLTVDKSRWQGNVFCSCVMHEALHNNHYTOKSLSPGK 228  
Db :|||||  
Db 300 SDGSFELYSLTVDKSRWQGNVFCSCVMHEALHNNHYTOKSLSPGK 347

RESULT 7  
US-08-459-657-43  
; Sequence 43, Application US/08459657  
; Patent No. 5914111  
; GENERAL INFORMATION:  
; APPLICANT: BIOGEN, INC.  
; APPLICANT: WALLNER, Barbara P.  
; APPLICANT: MILLER, Glenn T.  
; APPLICANT: ROSA, Margaret D.  
; TITLE OF INVENTION: CD2-BINDING DOMAIN OF LYMPHOCYTE  
; NUMBER OF SEQUENCES: 43  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Fish & Neave  
; CITY: New York  
; STATE: New York  
; COUNTRY: U.S.A.  
; ZIP: 10022-6250  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/459,657  
FILING DATE: 02-JUN-1995  
CLASSIFICATION: 424  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: PCT/US92/02050  
FILING DATE: 12-MAR-1992  
APPLICATION NUMBER: US 07/667,971  
FILING DATE: 12-MAR-1991  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/770,967  
FILING DATE: 07-OCT-1991  
ATTORNEY/AGENT INFORMATION:  
NAME: HALEY, James F., Jr.  
REGISTRATION NUMBER: 27,794  
REFERENCE/DOCKET NUMBER: B151CIP2  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (212)715-0600  
TELEFAX: (212)715-0673  
TELEX: 14-8367  
INFORMATION FOR SEQ ID NO: 43:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 347 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-459-657-43

Query Match 86.3%; Score 1234; DB 2; Length 347;  
Best Local Similarity 99.6%; Pred. No. 1.1e-39;  
Matches 227; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 MDKTHCTPCPCAPPELLGGPSVFLFPPKPKDTLMISRTPEVTCVVDVSHEDPEVKFNWYV 60  
Db 120 VDKTHCTPCPCAPPELLGGPSVFLFPPKPKDTLMISRTPEVTCVVDVSHEDPEVKFNWYV 179  
QY 61 DGEVHNATKPREOYNSTYRVSVTLVHQLDNLNGKEYKCKVSNKALPAPIETKISK 120  
Db 180 DGEVHNATKPREOYNSTYRVSVTLVHQLDNLNGKEYKCKVSNKALPAPIETKISK 239  
QY 121 KGQPREQVYTLPPSRDELTKNOVSLTCLVKGYPSPDIKAVESNGQPNNTKTPPVLD 180  
Db 240 KGQPREQVYTLPPSRDELTKNOVSLTCLVKGYPSPDIKAVESNGQPNNTKTPPVLD 299  
QY 181 SDGSFFLYSKLTVDKSRWQGNVFCSCVMHEALHNHYTKLSLSPGK 228  
Db 300 SDGSFFLYSKLTVDKSRWQGNVFCSCVMHEALHNHYTKLSLSPGK 347

RESULT 8  
US-08-460-132-43  
Sequence 43, Application US/08460132  
Patent No. 5928643  
GENERAL INFORMATION:  
APPLICANT: BIOGEN, INC.  
APPLICANT: WALLNER, Barbara P.  
APPLICANT: MILLER, Glenn T.  
APPLICANT: ROSA, Margaret D.  
TITLE OF INVENTION: CD2-BINDING DOMAIN OF LYMPHOCYTE  
NUMBER OF SEQUENCES: 43  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Fish & Neave  
STREET: 875 Third Avenue  
CITY: New York  
STATE: New York  
COUNTRY: U.S.A.  
ZIP: 10022-6250  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/460,132  
FILING DATE: 02-JUN-1995  
CLASSIFICATION: 424  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/940,861  
FILING DATE: 21-OCT-1992  
APPLICATION NUMBER: PCT/US92/02050  
FILING DATE: 12-MAR-1992  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/667,971  
FILING DATE: 12-MAR-1991  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/770,967  
FILING DATE: 07-OCT-1991  
ATTORNEY/AGENT INFORMATION:  
NAME: HALEY, James F., Jr.  
REGISTRATION NUMBER: 27,794  
REFERENCE/DOCKET NUMBER: B151CIP2  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (212)715-0600  
TELEFAX: (212)715-0673  
TELEX: 14-8367  
INFORMATION FOR SEQ ID NO: 43:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 347 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-460-132-43

Query Match 86.3%; Score 1234; DB 2; Length 347;  
Best Local Similarity 99.6%; Pred. No. 1.1e-39;  
Matches 227; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 MDKTHCTPCPCAPPELLGGPSVFLFPPKPKDTLMISRTPEVTCVVDVSHEDPEVKFNWYV 60  
Db 120 VDKTHCTPCPCAPPELLGGPSVFLFPPKPKDTLMISRTPEVTCVVDVSHEDPEVKFNWYV 179  
QY 61 DGEVHNATKPREOYNSTYRVSVTLVHQLDNLNGKEYKCKVSNKALPAPIETKISK 120  
Db 180 DGEVHNATKPREOYNSTYRVSVTLVHQLDNLNGKEYKCKVSNKALPAPIETKISK 239  
QY 121 KGQPREQVYTLPPSRDELTKNOVSLTCLVKGYPSPDIKAVESNGQPNNTKTPPVLD 180  
Db 240 KGQPREQVYTLPPSRDELTKNOVSLTCLVKGYPSPDIKAVESNGQPNNTKTPPVLD 299  
QY 181 SDGSFFLYSKLTVDKSRWQGNVFCSCVMHEALHNHYTKLSLSPGK 228  
Db 300 SDGSFFLYSKLTVDKSRWQGNVFCSCVMHEALHNHYTKLSLSPGK 347

RESULT 9  
US-08-466-465-8  
Sequence 8, Application US/08466465  
Patent No. 6162432  
GENERAL INFORMATION:  
APPLICANT: Wallner, Barbara P.  
APPLICANT: Cooper, Kevin D.  
TITLE OF INVENTION: Method of Prophylaxis or Treatment of Antigen  
PRESENTING CELL DRIVEN SKIN CONDITIONS USING  
INHIBITORS OF THE CD2/LFA-3 INTERACTION  
NUMBER OF SEQUENCES: 8  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: LAHIVE & COCKFIELD  
STREET: 60 State Street, Suite 510  
CITY: Boston  
STATE: Massachusetts  
COUNTRY: USA  
ZIP: 02109-1875

```
;
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/466,465
; FILING DATE:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/US92/08755
; FILING DATE: 06-OCT-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/862,022
; FILING DATE: 12-APR-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/770,969
; FILING DATE: 07-OCT-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: Myers, Louis (PLM)
; REGISTRATION NUMBER: 35,965
; REFERENCE/DOCKET NUMBER: BGP-111CP
; TELEPHONE: (617)227-7400
; TELEFAX: (617)227-5941
; INFORMATION FOR SEQ ID NO: 8:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 347 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-466-465-8

Query Match 86.3%; Score 1234; DB 4; Length 347;
Best Local Similarity 99.6%; Pred. No. 1.le-39;
Matches 227; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 MDKTHCTPCPAPPELLGSPSVFLPPPKDMLISRTPEVTCVVVDVSHEDPEVKFNWYV 60
Db :|||||
Db 120 VDKTHCTPCPAPPELLGSPSVFLPPPKDMLISRTPEVTCVVVDVSHEDPEVKFNWYV 179
QY 61 DGEVHNNAKTPREQYNTSYRVVSVLTVLHQDMLNGKEYCKVSNKALPAPIETKISK 120
Db :|||||
Db 180 DGEVHNNAKTPREQYNTSYRVVSVLTVLHQDMLNGKEYCKVSNKALPAPIETKISK 239
QY 121 KGQPREQVYTLPPSRDELTKNQVSLTCLVKGFYPSDIAVEWESNGQPENNYKTPPVLD 180
Db :|||||
Db 240 KGQPREQVYTLPPSRDELTKNQVSLTCLVKGFYPSDIAVEWESNGQPENNYKTPPVLD 299
QY 181 SDGSFFLYSKLTVDKSRWQGNVFSCSVMHAEALHNHYTQKSLSLSPGK 228
Db :|||||
Db 300 SDGSFFLYSKLTVDKSRWQGNVFSCSVMHAEALHNHYTQKSLSLSPGK 347

RESULT 10
PCT-US92-02050-43
; Sequence 43, Application PC/TUS9202050
; GENERAL INFORMATION:
; APPLICANT: BIOGEN, INC.
; APPLICANT: WALLNER, Barbara P.
; APPLICANT: MILLER, Glenn T.
; APPLICANT: ROSA, Margaret D.
; TITLE OF INVENTION: CD2-BINDING DOMAIN OF LYMPHOCYTE
; NUMBER OF SEQUENCES: 43
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Fish & Neave
; STREET: 875 Third Avenue
; CITY: New York
; STATE: New York
; COUNTRY: U.S.A.
; ZIP: 10022-6250
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
```

```
;
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US92/02050
; FILING DATE: 19920312
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/667,971
; FILING DATE: 12-MAR-1991
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/770,967
; FILING DATE: 07-OCT-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: HALEY, James F., Jr.
; REGISTRATION NUMBER: 27,794
; REFERENCE/DOCKET NUMBER: B15C1P2
; TELEPHONE: (212)715-0600
; TELEFAX: (212)715-0673
; TELEX: 14-8367
; INFORMATION FOR SEQ ID NO: 43:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 347 amino acids
; TYPE: AMINO ACID
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; PCT-US92-02050-43

Query Match 86.3%; Score 1234; DB 5; Length 347;
Best Local Similarity 99.6%; Pred. No. 1.le-39;
Matches 227; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 MDKTHCTPCPAPPELLGSPSVFLPPPKDMLISRTPEVTCVVVDVSHEDPEVKFNWYV 60
Db :|||||
Db 120 VDKTHCTPCPAPPELLGSPSVFLPPPKDMLISRTPEVTCVVVDVSHEDPEVKFNWYV 179
QY 61 DGEVHNNAKTPREQYNTSYRVVSVLTVLHQDMLNGKEYCKVSNKALPAPIETKISK 120
Db :|||||
Db 180 DGEVHNNAKTPREQYNTSYRVVSVLTVLHQDMLNGKEYCKVSNKALPAPIETKISK 239
QY 121 KGQPREQVYTLPPSRDELTKNQVSLTCLVKGFYPSDIAVEWESNGQPENNYKTPPVLD 180
Db :|||||
Db 240 KGQPREQVYTLPPSRDELTKNQVSLTCLVKGFYPSDIAVEWESNGQPENNYKTPPVLD 299
QY 181 SDGSFFLYSKLTVDKSRWQGNVFSCSVMHAEALHNHYTQKSLSLSPGK 228
Db :|||||
Db 300 SDGSFFLYSKLTVDKSRWQGNVFSCSVMHAEALHNHYTQKSLSLSPGK 347

RESULT 11
US-09-189-129-2
; Sequence 2, Application US/09189129
; Patent No. 6323027
; GENERAL INFORMATION:
; APPLICANT: Burkly, Linda C
; APPLICANT: Benjamin, Christopher D
; APPLICANT: Hession, Catherine A
; APPLICANT: Whitty, Adrian
; TITLE OF INVENTION: COMMON GAMMA CHAIN BLOCKING AGENTS
; NUMBER OF SEQUENCES: 17
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Biogen, Inc.
; STREET: 14 Cambridge Center
; CITY: Cambridge
; STATE: Massachusetts
; COUNTRY: USA
; ZIP: 02142
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
```

;; CURRENT APPLICATION DATA:  
;; APPLICATION NUMBER: US/09/189,129  
;; FILING DATE:  
;; PRIOR APPLICATION DATA:  
;; APPLICATION NUMBER: A006 PCT CIP  
;; FILING DATE: 09-MAY-1997  
;; APPLICATION NUMBER: 60/017,466  
;; FILING DATE: 10-MAY-1996  
;; ATTORNEY/AGENT INFORMATION:  
;; NAME: Kaplan, Warren A.  
;; REGISTRATION NUMBER: 34,199  
;; REFERENCE/DOCKET NUMBER: A006 PCT CIP  
;; TELECOMMUNICATION INFORMATION:  
;; TELEPHONE: 617 679-2000  
;; TELEFAX: 617 679-2838  
;; INFORMATION FOR SEQ ID NO: 2:  
;; SEQUENCE CHARACTERISTICS:  
;; LENGTH: 482 amino acids  
;; TYPE: amino acid  
;; STRANDEDNESS: single  
;; TOPOLOGY: linear  
;; MOLECULE TYPE: YES  
;; HYPOTHETICAL: NO  
;; ANTI-SENSE: NO  
US-09-189-129-2

Query Match 86.3%; Score 1234; DB 4; Length 482;  
Best Local Similarity 99.6%; Pred. No. 2.4e-39;  
Matches 227; Conservative 1; Mismatches 0; Indels 0; Gaps 0;  
QY 1 MKDTHCTPCPCAPPELLGGPSVFLFPPKPKDTLMISRTPEVTCVVDVSHEDPEVKFNWYV 60  
Db 255 VDKTHCTPCPCAPPELLGGPSVFLFPPKPKDTLMISRTPEVTCVVDVSHEDPEVKFNWYV 314  
QY 61 DGEVHNKAKTPREEQYNSTYRVVSVLTVLHQDWLNGKEYKCKVSNKALPAPIEKTISKAK 120  
Db 315 DGEVHNKAKTPREEQYNSTYRVVSVLTVLHQDWLNGKEYKCKVSNKALPAPIEKTISKAK 374  
QY 121 KGPREPQVYTLPPSRDELTKNOVSLTCLVKGYFSPDIAVEWESNGOPENNYKTTTPVLD 180  
Db 375 KGPREPQVYTLPPSRDELTKNOVSLTCLVKGYFSPDIAVEWESNGOPENNYKTTTPVLD 434  
QY 181 SDGSFFLYSKLTVDKSRWQGNVFSCSVMHEALHNHYTQKSLSLSPGK 228  
Db 435 SDGSFFLYSKLTVDKSRWQGNVFSCSVMHEALHNHYTQKSLSLSPGK 482

RESULT 12  
US-08-595-043A-50  
; Sequence 50, Application US/08595043A  
; Patent No. 5935824  
; GENERAL INFORMATION:  
; APPLICANT: SGARLATO, GREGORY D.  
; TITLE OF INVENTION: PROTEIN EXPRESSION SYSTEM  
; NUMBER OF SEQUENCES: 90  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: MEDLEN & CARROLL  
; STREET: 220 MONTGOMERY STREET, SUITE 2200  
; CITY: SAN FRANCISCO  
; STATE: CALIFORNIA  
; COUNTRY: UNITED STATES OF AMERICA  
; ZIP: 94104  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent in Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/595,043A  
; FILING DATE: 31-JAN-1996  
; CLASSIFICATION: 435  
; ATTORNEY/AGENT INFORMATION:  
; NAME: CARROLL, PETER G.

;; REGISTRATION NUMBER: 32,837  
;; REFERENCE/DOCKET NUMBER: SGAR-00371  
;; TELECOMMUNICATION INFORMATION:  
;; TELEPHONE: (415) 705-8410  
;; TELEFAX: (415) 397-8338  
;; INFORMATION FOR SEQ ID NO: 50:  
;; SEQUENCE CHARACTERISTICS:  
;; LENGTH: 232 amino acids  
;; TYPE: amino acid  
;; TOPOLOGY: linear  
;; MOLECULE TYPE: protein  
US-08-595-043A-50

Query Match 86.2%; Score 1233; DB 2; Length 232;  
Best Local Similarity 100.0%; Pred. No. 4.5e-40;  
Matches 227; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 2 DKTHCTPCPCAPPELLGGPSVFLFPPKPKDTLMISRTPEVTCVVDVSHEDPEVKFNWYVD 61  
Db 6 DKTHCTPCPCAPPELLGGPSVFLFPPKPKDTLMISRTPEVTCVVDVSHEDPEVKFNWYVD 65  
QY 62 GVEVHNKAKTPREEQYNSTYRVVSVLTVLHQDWLNGKEYKCKVSNKALPAPIEKTISKAK 121  
Db 66 GVEVHNKAKTPREEQYNSTYRVVSVLTVLHQDWLNGKEYKCKVSNKALPAPIEKTISKAK 125  
QY 122 GOPREPOVYTLPPSRDELTKNOVSLTCLVKGYFSPDIAVEWESNGOPENNYKTTTPVLD 181  
Db 126 GOPREPOVYTLPPSRDELTKNOVSLTCLVKGYFSPDIAVEWESNGOPENNYKTTTPVLD 185  
QY 182 DGSFFLYSKLTVDKSRWQGNVFSCSVMHEALHNHYTQKSLSLSPGK 228  
Db 186 DGSFFLYSKLTVDKSRWQGNVFSCSVMHEALHNHYTQKSLSLSPGK 232

RESULT 13  
US-09-131-247-6  
; Sequence 6, Application US/09131247  
; Patent No. 6294170  
; GENERAL INFORMATION:  
; APPLICANT: Boone, Thomas C.  
; APPLICANT: Hershenson, Susan  
; APPLICANT: Bevilacqua, Michael P.  
; APPLICANT: Collins, David S.  
; TITLE OF INVENTION: COMPOSITION AND METHOD FOR TREATING INFLAMMATORY  
; TITLE OF INVENTION: DISEASES  
; FILE REFERENCE: A-365F.  
; CURRENT APPLICATION NUMBER: US/09/131,247  
; CURRENT FILING DATE: 1998-08-07  
; EARLIER APPLICATION NUMBER: 60/055,185  
; EARLIER FILING DATE: 1997-08-08  
; EARLIER APPLICATION NUMBER: PCT/US 97/02131  
; EARLIER FILING DATE: 1997-02-10  
; NUMBER OF SEQ ID NOS: 16  
; SOFTWARE: Patent in Ver. 2.0  
; SEQ ID NO 6  
; LENGTH: 235  
; TYPE: PRT  
; ORGANISM: Human  
US-09-131-247-6

Query Match 86.2%; Score 1233; DB 4; Length 235;  
Best Local Similarity 100.0%; Pred. No. 4.7e-40;  
Matches 227; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 2 DKTHCTPCPCAPPELLGGPSVFLFPPKPKDTLMISRTPEVTCVVDVSHEDPEVKFNWYVD 61  
Db 9 DKTHCTPCPCAPPELLGGPSVFLFPPKPKDTLMISRTPEVTCVVDVSHEDPEVKFNWYVD 68  
QY 62 GVEVHNKAKTPREEQYNSTYRVVSVLTVLHQDWLNGKEYKCKVSNKALPAPIEKTISKAK 121  
Db 69 GVEVHNKAKTPREEQYNSTYRVVSVLTVLHQDWLNGKEYKCKVSNKALPAPIEKTISKAK 128  
QY 122 GOPREPOVYTLPPSRDELTKNOVSLTCLVKGYFSPDIAVEWESNGOPENNYKTTTPVLD 181

Db 129 GQPREPQVYTLPPSRDELTKNQVSLTCLVKGFYPSDIAVEWESNGQPENNYKTTTPPVLDL 188  
QY 182 DGSFFLYSKLTVDKSRWQGNVFSCVMEALHNHYTQKSLSLSPGK 228  
Db 189 DGSFFLYSKLTVDKSRWQGNVFSCVMEALHNHYTQKSLSLSPGK 235

RESULT 14  
US-09-178-869-2  
; Sequence 2, Application US/09178869B  
; Patent No. 6197294  
; GENERAL INFORMATION:  
; APPLICANT: Wong, Shou  
; APPLICANT: Hickey, William F.  
; APPLICANT: Hamman, Joseph P.  
; APPLICANT: Baetge, E. Edward  
; TITLE OF INVENTION: CELL SURFACE-INDUCED MACROPHAGE ACTIVATION  
; FILE REFERENCE: 17810-043  
; CURRENT APPLICATION NUMBER: US/09/178,869B  
; CURRENT FILING DATE: 1998-10-26  
; NUMBER OF SEQ ID NOS: 14  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 2  
; LENGTH: 331  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-09-178-869-2

Query Match 86.2%; Score 1233; DB 4; Length 331;  
Best Local Similarity 100.0%; Pred. No. 1.1e-39;  
Matches 227; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 2 DKTHCTCPAPPELLGGPSVFLFPPPKDRLMISRTPEVTCVVVDVSHEDPEVKFNWYVD 61  
Db 105 DKTHCTCPAPPELLGGPSVFLFPPPKDRLMISRTPEVTCVVVDVSHEDPEVKFNWYVD 164  
QY 62 GVEVHNATKPREEQYNSTYRVVSVLTVLHODWLNKGEYKCKVSNKALPAPIEKTISKAK 121  
Db 165 GVEVHNATKPREEQYNSTYRVVSVLTVLHODWLNKGEYKCKVSNKALPAPIEKTISKAK 224  
QY 122 GQPREPQVYTLPPSRDELTKNQVSLTCLVKGFYPSDIAVEWESNGQPENNYKTTTPPVLDL 181  
Db 225 GQPREPQVYTLPPSRDELTKNQVSLTCLVKGFYPSDIAVEWESNGQPENNYKTTTPPVLDL 284  
QY 182 DGSFFLYSKLTVDKSRWQGNVFSCVMEALHNHYTQKSLSLSPGK 228  
Db 285 DGSFFLYSKLTVDKSRWQGNVFSCVMEALHNHYTQKSLSLSPGK 331

RESULT 15  
US-09-180-100-11  
; Sequence 11, Application US/09180100  
; Patent No. 6306395  
; GENERAL INFORMATION:  
; APPLICANT: NAKAMURA, No. 630639510  
; APPLICANT: NAGATA, Shigekazu  
; TITLE OF INVENTION: NOVEL FAS ANTIGEN DERIVATIVE  
; FILE REFERENCE: 1110-207P  
; CURRENT APPLICATION NUMBER: US/09/180,100  
; CURRENT FILING DATE: 1998-11-02  
; EARLIER APPLICATION NUMBER: PCT/JP97/01502  
; EARLIER FILING DATE: 1997-05-01  
; NUMBER OF SEQ ID NOS: 25  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 11  
; LENGTH: 360  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-09-180-100-11

Query Match 86.2%; Score 1233; DB 4; Length 360;

Best Local Similarity 100.0%; Pred. No. 1.3e-39;  
Matches 227; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 2 DKTHCTCPAPPELLGGPSVFLFPPPKDRLMISRTPEVTCVVVDVSHEDPEVKFNWYVD 61  
Db 134 DKTHCTCPAPPELLGGPSVFLFPPPKDRLMISRTPEVTCVVVDVSHEDPEVKFNWYVD 193  
QY 62 GVEVHNATKPREEQYNSTYRVVSVLTVLHODWLNKGEYKCKVSNKALPAPIEKTISKAK 121  
Db 194 GVEVHNATKPREEQYNSTYRVVSVLTVLHODWLNKGEYKCKVSNKALPAPIEKTISKAK 253  
QY 122 GQPREPQVYTLPPSRDELTKNQVSLTCLVKGFYPSDIAVEWESNGQPENNYKTTTPPVLDL 181  
Db 254 GQPREPQVYTLPPSRDELTKNQVSLTCLVKGFYPSDIAVEWESNGQPENNYKTTTPPVLDL 313  
QY 182 DGSFFLYSKLTVDKSRWQGNVFSCVMEALHNHYTQKSLSLSPGK 228  
Db 314 DGSFFLYSKLTVDKSRWQGNVFSCVMEALHNHYTQKSLSLSPGK 360

Search completed: April 21, 2003, 10:50:27  
Job time : 16.7529 secs

GenCore version 5.1.4.p5.4578  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: April 21, 2003, 10:43:21 ; Search time 12.1647 Seconds  
(without alignments)  
1640.982 Million cell updates/sec

Title: 2LINK7LINK7  
Perfect score: 1430  
Sequence: 1 MDKTHCPCPAPPELLGGPS.....AFDNLKAFYDKVAEKLKEAF 264

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.0

Searched: 288829 seqs, 7561385 residues

Total number of hits satisfying chosen parameters: 288829

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published\_Applications\_AA:\*

- 1: /cgn2\_6/ptodata/2/pubpaa/US08\_NEW\_PUB.pap:\*
- 2: /cgn2\_6/ptodata/2/pubpaa/PCT\_NEW\_PUB.pap:\*
- 3: /cgn2\_6/ptodata/2/pubpaa/US06\_NEW\_PUB.pap:\*
- 4: /cgn2\_6/ptodata/2/pubpaa/US06\_PUBCOMB.pap:\*
- 5: /cgn2\_6/ptodata/2/pubpaa/US07\_NEW\_PUB.pap:\*
- 6: /cgn2\_6/ptodata/2/pubpaa/US07\_PUBCOMB.pap:\*
- 7: /cgn2\_6/ptodata/2/pubpaa/US08\_PUBCOMB.pap:\*
- 8: /cgn2\_6/ptodata/2/pubpaa/US08\_PUBCOMB.pap:\*
- 9: /cgn2\_6/ptodata/2/pubpaa/US09\_NEW\_PUB.pap:\*
- 10: /cgn2\_6/ptodata/2/pubpaa/US09\_PUBCOMB.pap:\*
- 11: /cgn2\_6/ptodata/2/pubpaa/US10\_NEW\_PUB.pap:\*
- 12: /cgn2\_6/ptodata/2/pubpaa/US10\_PUBCOMB.pap:\*
- 13: /cgn2\_6/ptodata/2/pubpaa/US60\_NEW\_PUB.pap:\*
- 14: /cgn2\_6/ptodata/2/pubpaa/US60\_PUBCOMB.pap:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1260	88.1	729	10	US-09-825-012-52
2	1260	88.1	739	10	US-09-825-012-61
3	1257	87.9	731	10	US-09-825-012-46
4	1257	87.9	741	10	US-09-825-012-55
5	1250	87.4	730	10	US-09-825-012-49
6	1250	87.4	740	10	US-09-825-012-58
7	1240	86.7	389	10	US-09-784-623-14
8	1238	86.6	228	9	US-09-847-277-2
9	1238	86.6	228	9	US-09-847-249A-2
10	1238	86.6	228	9	US-09-840-669B-2
11	1238	86.6	228	9	US-09-843-221A-2
12	1238	86.6	228	10	US-09-847-712-2
13	1238	86.6	282	9	US-09-840-277-109
14	1237	86.5	394	10	US-09-854-864-31
15	1234	86.3	229	9	US-10-215-297-2
16	1234	86.3	229	9	US-10-215-298-2
17	1234	86.3	347	9	US-10-091-236-17
18	1234	86.3	347	9	US-10-091-313-7
19	1234	86.3	347	9	US-10-091-268-7

20	1234	86.3	347	10	US-09-796-033-8	Sequence 8, Appli
21	1234	86.3	347	10	US-09-730-465-8	Sequence 8, Appli
22	1234	86.3	380	10	US-09-948-018-36	Sequence 36, Appli
23	1234	86.3	380	10	US-09-948-018-39	Sequence 39, Appli
24	1234	86.3	388	10	US-09-734-300-8	Sequence 8, Appli
25	1234	86.3	388	10	US-09-734-300-9	Sequence 9, Appli
26	1234	86.3	399	9	US-09-832-659-2	Sequence 2, Appli
27	1234	86.3	418	9	US-09-832-659-42	Sequence 42, Appli
28	1234	86.3	423	9	US-09-832-659-44	Sequence 44, Appli
29	1234	86.3	482	10	US-09-824-286-2	Sequence 2, Appli
30	1234	86.3	594	10	US-09-815-108-22	Sequence 22, Appli
31	1233	86.2	232	10	US-09-996-357-10	Sequence 10, Appli
32	1233	86.2	235	10	US-09-784-623-6	Sequence 6, Appli
33	1233	86.2	247	10	US-09-996-357-13	Sequence 13, Appli
34	1233	86.2	267	10	US-09-996-357-12	Sequence 12, Appli
35	1233	86.2	281	10	US-09-854-864-10	Sequence 10, Appli
36	1233	86.2	283	10	US-09-854-864-9	Sequence 9, Appli
37	1233	86.2	330	9	US-10-047-542-20	Sequence 20, Appli
38	1233	86.2	330	9	US-09-995-898A-15	Sequence 15, Appli
39	1233	86.2	360	10	US-09-949-713-11	Sequence 11, Appli
40	1233	86.2	376	10	US-09-949-713-22	Sequence 22, Appli
41	1233	86.2	388	10	US-09-784-623-16	Sequence 16, Appli
42	1233	86.2	397	10	US-09-854-864-18	Sequence 18, Appli
43	1233	86.2	404	10	US-09-948-018-16	Sequence 16, Appli
44	1233	86.2	443	9	US-10-166-232A-5	Sequence 5, Appli
45	1233	86.2	451	9	US-10-077-023-17	Sequence 17, Appli

ALIGNMENTS

RESULT 1  
US-09-825-012-52  
; Sequence 52, Application US/09825012  
; Patent No. US2002012798A1  
; GENERAL INFORMATION:  
; APPLICANT: Young, Robert  
; TITLE OF INVENTION: Compounds for Targeting  
; FILE REFERENCE: 43191-256808  
; CURRENT APPLICATION NUMBER: US/09/825,012  
; CURRENT FILING DATE: 2001-04-03  
; PRIOR APPLICATION NUMBER: US 60/237,159  
; PRIOR FILING DATE: 2000-10-02  
; PRIOR APPLICATION NUMBER: GB 0008049.9  
; PRIOR FILING DATE: 2000-04-03  
; NUMBER OF SEQ ID NOS: 102  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 52  
; LENGTH: 729  
; TYPE: PRT  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Humanised HMFG1 heavy chain - DNase I fusion  
US-09-825-012-52

Query Match	88.1%	Score	1260;	DB	10;	Length	729;
Best Local Similarity	50.5%	Pred	No. 1.6e-31;				
Matches	241;	Conservative	6;	Mismatches	15;	Indels	215;
Gaps	5;						
QY	2	DKHTCPCPAPPELLGGPSVFLFPPKPKDTLMISRTPEVTCVVDVSHEDPEVKFNMYVD	61				
Db	241	DKHTCPCPAPPELLGGPSVFLFPPKPKDTLMISRTPEVTCVVDVSHEDPEVKFNMYVD	300				
QY	62	GVEVHNKTPREPOYNSTRVWSVLVFLHODWLNKGYKCKVSNKALPAPIETISKAK	121				
Db	301	GVEVHNKTPREPOYNSTRVWSVLVFLHODWLNKGYKCKVSNKALPAPIETISKAK	360				
QY	122	GOPREPOVYTLPPSRDELTKNOVSLTCLVKGFPDSDAVESNGQPNNTKTPPVLD	181				
Db	361	GOPREPOVYTLPPSRDELTKNOVSLTCLVKGFPDSDAVESNGQPNNTKTPPVLD	420				
QY	182	DGSPFLYKSLTVDKSRQOQGNVFCVSNHAEALHNYHTOKLSLSLSPGKDWLK-----	232				

Db 421 DGSFFLYSKLTVDKSRWQGNVFCSCVMHEALHNHYTKQSLSPGSGGLKIAAFNIQTF 480  
QY 233 -----AFYD---KVAEK----- 232  
Db 481 GETKMSNATLVSYIVQILSRDYIALVQEVDRDLSHTAVGKLLDNLNODAPDTHYVYVSEPL 540  
QY 233 ----- 232  
Db 541 GRNSYKERYLVYRDPQVSAVDSYDDGCEPCGNDTFNREPAIVRFFSRFTEVREFAIV 600  
QY 233 -----AYD---KVAEK-----LKEA 245  
Db 601 PLHAAPGDAVAEIDALYDVLVDQEKWGLEDMVLMGDFNAGCSYVRPSQWSSIRLWTSPT 660  
QY 246 EDWL-----KAFYDKV-----AEKLEA 263  
Db 661 FQWLIPDSADTTATPTHCAIDRIVVAGMLLRGAVPDSALPFNFQAAAYGLSDOLAQA 717  
RESULT 2  
US-09-825-012-61  
; Sequence 61, Application US/09825012  
; Patent No. US20020122798A1  
; GENERAL INFORMATION:  
; APPLICANT: Young, Robert  
; TITLE OF INVENTION: Compounds for Targeting  
; FILE REFERENCE: 43191-256808  
; CURRENT APPLICATION NUMBER: US/09/825,012  
; PRIOR FILING DATE: 2001-04-03  
; PRIOR APPLICATION NUMBER: US 60/237,159  
; PRIOR FILING DATE: 2000-10-02  
; PRIOR APPLICATION NUMBER: GB 0008049.9  
; PRIOR FILING DATE: 2000-04-03  
; NUMBER OF SEQ ID NOS: 102  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 61  
; LENGTH: 739  
; TYPE: PRT  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Humanised HMF1 heavy chain - DNase I fusion  
US-09-825-012-61  
Query Match 88.1%; Score 1260; DB 10; Length 739;  
Best Local Similarity 50.5%; Pred. No. 1.6e-31;  
Matches 241; Conservative 6; Mismatches 15; Indels 215; Gaps 5;  
QY 2 DKHTCCPCPAPPELLGGPSVFLFPPKPKDTLMISRTPEVTCVVVDVSHEDPEVKFNWYVD 61  
Db 241 DKHTCCPCPAPPELLGGPSVFLFPPKPKDTLMISRTPEVTCVVVDVSHEDPEVKFNWYVD 300  
QY 62 GVEVHNATKPREQYNSTYRVVSVLTVLHQDWLNGKEYCKVSNKALPAPIETISKAK 121  
Db 301 GVEVHNATKPREQYNSTYRVVSVLTVLHQDWLNGKEYCKVSNKALPAPIETISKAK 360  
QY 122 GQPREPQVYTLPPSRDELTKNQVSLTCLVKGFYPSDIAVEWESNGQPENNYKTTTPPVLD 181  
Db 361 GQPREPQVYTLPPSRDELTKNQVSLTCLVKGFYPSDIAVEWESNGQPENNYKTTTPPVLD 420  
QY 182 DGSFFLYSKLTVDKSRWQGNVFCSCVMHEALHNHYTKQSLSPGSGGLKIAAFNIQTF 480  
Db 421 DGSFFLYSKLTVDKSRWQGNVFCSCVMHEALHNHYTKQSLSPGSGGLKIAAFNIQTF 540  
QY 233 -----AFYD---KVAEK----- 232  
Db 481 GETKMSNATLVSYIVQILSRDYIALVQEVDRDLSHTAVGKLLDNLNODAPDTHYVYVSEPL 540  
QY 233 ----- 232  
Db 541 GRNSYKERYLVYRDPQVSAVDSYDDGCEPCGNDTFNREPAIVRFFSRFTEVREFAIV 600  
QY 233 -----AFYD---KVAEK-----LKEA 245

Db 601 PLHAAPGDAVAEIDALYDVLVDQEKWGLEDMVLMGDFNAGCSYVRPSQWSSIRLWTSPT 660  
QY 246 EDWL-----KAFYDKV-----AEKLEA 263  
Db 661 FQWLIPDSADTTATPTHCAIDRIVVAGMLLRGAVPDSALPFNFQAAAYGLSDOLAQA 717  
RESULT 3  
US-09-825-012-46  
; Sequence 46, Application US/09825012  
; Patent No. US20020122798A1  
; GENERAL INFORMATION:  
; APPLICANT: Young, Robert  
; TITLE OF INVENTION: Compounds for Targeting  
; FILE REFERENCE: 43191-256808  
; CURRENT APPLICATION NUMBER: US/09/825,012  
; PRIOR FILING DATE: 2001-04-03  
; PRIOR APPLICATION NUMBER: US 60/237,159  
; PRIOR FILING DATE: 2000-10-02  
; PRIOR APPLICATION NUMBER: GB 0008049.9  
; PRIOR FILING DATE: 2000-04-03  
; NUMBER OF SEQ ID NOS: 102  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 46  
; LENGTH: 731  
; TYPE: PRT  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Humanised HMF1 heavy chain - DNase I fusion  
US-09-825-012-46  
Query Match 87.9%; Score 1257; DB 10; Length 731;  
Best Local Similarity 50.1%; Pred. No. 2e-31; Mismatches 6; Indels 217; Gaps 4;  
Matches 240; Conservative 6; Mismatches 6; Indels 217; Gaps 4;  
QY 2 DKHTCCPCPAPPELLGGPSVFLFPPKPKDTLMISRTPEVTCVVVDVSHEDPEVKFNWYVD 61  
Db 241 DKHTCCPCPAPPELLGGPSVFLFPPKPKDTLMISRTPEVTCVVVDVSHEDPEVKFNWYVD 300  
QY 62 GVEVHNATKPREQYNSTYRVVSVLTVLHQDWLNGKEYCKVSNKALPAPIETISKAK 121  
Db 301 GVEVHNATKPREQYNSTYRVVSVLTVLHQDWLNGKEYCKVSNKALPAPIETISKAK 360  
QY 122 GQPREPQVYTLPPSRDELTKNQVSLTCLVKGFYPSDIAVEWESNGQPENNYKTTTPPVLD 181  
Db 361 GQPREPQVYTLPPSRDELTKNQVSLTCLVKGFYPSDIAVEWESNGQPENNYKTTTPPVLD 420  
QY 182 DGSFFLYSKLTVDKSRWQGNVFCSCVMHEALHNHYTKQSLSPGSGGLKIAAFNIQ 480  
Db 421 DGSFFLYSKLTVDKSRWQGNVFCSCVMHEALHNHYTKQSLSPGSGGLKIAAFNIQ 480  
QY 229 ----- 228  
Db 481 TFGETKMSNATLVSYIVQILSRDYIALVQEVDRDLSHTAVGKLLDNLNODAPDTHYVYVSE 540  
QY 229 ----- 228  
Db 541 PLGRNSYKERYLVYRDPQVSAVDSYDDGCEPCGNDTFNREPAIVRFFSRFTEVREFA 600  
QY 229 -----DWLKAFYDKVAK-----LK 243  
Db 601 IVPLHAAPGDAVAEIDALYDVLVDQEKWGLEDMVLMGDFNAGCSYVRPSQWSSIRLWTS 660  
QY 244 EAFDWL-----KAFYDKV-----AEKLEA 263  
Db 661 PTFQWLIPDSADTTATPTHCAIDRIVVAGMLLRGAVPDSALPFNFQAAAYGLSDOLAQA 719  
RESULT 4  
US-09-825-012-55  
; Sequence 55, Application US/09825012  
; Patent No. US20020122798A1  
; GENERAL INFORMATION:



```
; APPLICANT: Young, Robert
; TITLE OF INVENTION: Compounds for Targeting
; FILE REFERENCE: 43191-256808
; CURRENT APPLICATION NUMBER: US/09/825,012
; CURRENT FILING DATE: 2001-04-03
; PRIOR APPLICATION NUMBER: US 60/237,159
; PRIOR FILING DATE: 2000-10-02
; PRIOR APPLICATION NUMBER: GB 0008049.9
; PRIOR FILING DATE: 2000-04-03
; NUMBER OF SEQ ID NOS: 102
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 55
; LENGTH: 741
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Humanised HMFG1 heavy chain - DNase I fusion
US-09-825-012-55

Query Match      87.9%; Score 1257; DB 10; Length 741;
Best Local Similarity 50.1%; Pred. No. 2e-31;
Matches 240; Conservative 6; Mismatches 16; Indels 217; Gaps 4;

QY  2 DKHTCPCPAPELLGGPSVFLFPPKPKDTLMISRTPEVTCVVDVSHEDPEVKFNWYVD 61
Db  241 DKHTCPCPAPELLGGPSVFLFPPKPKDTLMISRTPEVTCVVDVSHEDPEVKFNWYVD 300
QY  62 GVEVHNKTKPREEQYNSTYRVVSVLTVLHODWLNQGEYCKCKVSNKALPAPIETISKAK 121
Db  301 GVEVHNKTKPREEQYNSTYRVVSVLTVLHODWLNQGEYCKCKVSNKALPAPIETISKAK 360
QY  122 GQPREPQVYTLPPSRDELTKNQVSLTCLVKGFYPSDIAVEWESNGQPENNYKTTTPPVLD 181
Db  361 GQPREPQVYTLPPSRDELTKNQVSLTCLVKGFYPSDIAVEWESNGQPENNYKTTTPPVLD 420
QY  182 DGSFPLYSKLTVDKSRWQQGNVFSCSVMHEALHNHYTQKSLSLSPGK----- 228
Db  421 DGSFPLYSKLTVDKSRWQQGNVFSCSVMHEALHNHYTQKSLSLSPGK----- 480
QY  229 -----KAFYDKV----- 228
Db  481 TFGETKMSNATLVSYIVQILSRDYALVQEVDRSHLTAVGKLLDNLNODAPDTYHYVYVSE 540
QY  229 ----- 228
Db  541 PLGRNSYKERYLFVYRPDQVSAVDYDDGCEPCGNDTFNREPAIVRFFSRTFVREFA 600
QY  229 -----DWLKAFYDKVAEK-----LK 243
Db  601 IVPLHAAPGDAVAEIDALYDVLVDVQEKWGLDVMGMDFNAGCSYVRPQWSSIRLWTS 660
QY  244 EAFDWL-----KAFYDKV-----AEKLKEA 263
Db  661 PTFOWLIPDSADTTATPTTHCAYDRIVVAGMLLRGAVVPDSALPFPNFOAAYGLSDQLAQA 719

RESULT 5
US-09-825-012-49
; Sequence 49, Application US/09825012
; Patent No. US200201227981
; GENERAL INFORMATION:
; APPLICANT: Young, Robert
; TITLE OF INVENTION: Compounds for Targeting
; FILE REFERENCE: 43191-256808
; CURRENT APPLICATION NUMBER: US/09/825,012
; CURRENT FILING DATE: 2001-04-03
; PRIOR APPLICATION NUMBER: US 60/237,159
; PRIOR FILING DATE: 2000-10-02
; PRIOR APPLICATION NUMBER: GB 0008049.9
; PRIOR FILING DATE: 2000-04-03
; NUMBER OF SEQ ID NOS: 102
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 49

Query Match      87.9%; Score 1257; DB 10; Length 740;
Best Local Similarity 50.4%; Pred. No. 3.4e-31;
Matches 241; Conservative 6; Mismatches 15; Indels 216; Gaps 6;

QY  2 DKHTCPCPAPELLGGPSVFLFPPKPKDTLMISRTPEVTCVVDVSHEDPEVKFNWYVD 61
Db  241 DKHTCPCPAPELLGGPSVFLFPPKPKDTLMISRTPEVTCVVDVSHEDPEVKFNWYVD 300
QY  62 GVEVHNKTKPREEQYNSTYRVVSVLTVLHODWLNQGEYCKCKVSNKALPAPIETISKAK 121
Db  301 GVEVHNKTKPREEQYNSTYRVVSVLTVLHODWLNQGEYCKCKVSNKALPAPIETISKAK 360
QY  122 GQPREPQVYTLPPSRDELTKNQVSLTCLVKGFYPSDIAVEWESNGQPENNYKTTTPPVLD 181
Db  361 GQPREPQVYTLPPSRDELTKNQVSLTCLVKGFYPSDIAVEWESNGQPENNYKTTTPPVLD 420
QY  182 DGSFPLYSKLTVDKSRWQQGNVFSCSVMHEALHNHYTQKSLSLSPGK----- 232
Db  421 DGSFPLYSKLTVDKSRWQQGNVFSCSVMHEALHNHYTQKSLSLSPGK----- 480
QY  233 ----- 232
Db  481 FGETKMSNATLVSYIVQILSRDYALVQEVDRSHLTAVGKLLDNLNODAPDTYHYVYVSE 540
QY  233 ----- 232
Db  541 LGRNSYKERYLFVYRPDQVSAVDYDDGCEPCGNDTFNREPAIVRFFSRTFVREFAI 600
QY  233 -----AFYD---KVAEK-----LKE 244
Db  601 VPLHAAPGDAVAEIDALYDVLVDVQEKWGLDVMGMDFNAGCSYVRPQWSSIRLWTS 660
QY  245 AFDWL-----KAFYDKV-----AEKLKEA 263
Db  661 TFWLIPDSADTTATPTTHCAYDRIVVAGMLLRGAVVPDSALPFPNFOAAYGLSDQLAQA 718

RESULT 6
US-09-825-012-58
; Sequence 58, Application US/09825012
; Patent No. US200201227981
; GENERAL INFORMATION:
; APPLICANT: Young, Robert
; TITLE OF INVENTION: Compounds for Targeting
; FILE REFERENCE: 43191-256808
; CURRENT APPLICATION NUMBER: US/09/825,012
; CURRENT FILING DATE: 2001-04-03
; PRIOR APPLICATION NUMBER: US 60/237,159
; PRIOR FILING DATE: 2000-10-02
; PRIOR APPLICATION NUMBER: GB 0008049.9
; PRIOR FILING DATE: 2000-04-03
; NUMBER OF SEQ ID NOS: 102
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 58
; LENGTH: 740
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Humanised HMFG1 heavy chain - DNase I fusion
US-09-825-012-58

Query Match      87.4%; Score 1250; DB 10; Length 740;
Best Local Similarity 50.4%; Pred. No. 3.4e-31;
Matches 241; Conservative 6; Mismatches 15; Indels 216; Gaps 6;

QY  2 DKHTCPCPAPELLGGPSVFLFPPKPKDTLMISRTPEVTCVVDVSHEDPEVKFNWYVD 61
Db  241 DKHTCPCPAPELLGGPSVFLFPPKPKDTLMISRTPEVTCVVDVSHEDPEVKFNWYVD 300
QY  62 GVEVHNKTKPREEQYNSTYRVVSVLTVLHODWLNQGEYCKCKVSNKALPAPIETISKAK 121
Db  301 GVEVHNKTKPREEQYNSTYRVVSVLTVLHODWLNQGEYCKCKVSNKALPAPIETISKAK 360
QY  122 GQPREPQVYTLPPSRDELTKNQVSLTCLVKGFYPSDIAVEWESNGQPENNYKTTTPPVLD 181
Db  361 GQPREPQVYTLPPSRDELTKNQVSLTCLVKGFYPSDIAVEWESNGQPENNYKTTTPPVLD 420
QY  182 DGSFPLYSKLTVDKSRWQQGNVFSCSVMHEALHNHYTQKSLSLSPGK----- 232
Db  421 DGSFPLYSKLTVDKSRWQQGNVFSCSVMHEALHNHYTQKSLSLSPGK----- 480
QY  233 ----- 232
Db  481 FGETKMSNATLVSYIVQILSRDYALVQEVDRSHLTAVGKLLDNLNODAPDTYHYVYVSE 540
QY  233 ----- 232
Db  541 LGRNSYKERYLFVYRPDQVSAVDYDDGCEPCGNDTFNREPAIVRFFSRTFVREFAI 600
QY  233 -----AFYD---KVAEK-----LKE 244
Db  601 VPLHAAPGDAVAEIDALYDVLVDVQEKWGLDVMGMDFNAGCSYVRPQWSSIRLWTS 660
QY  245 AFDWL-----KAFYDKV-----AEKLKEA 263
Db  661 TFWLIPDSADTTATPTTHCAYDRIVVAGMLLRGAVVPDSALPFPNFOAAYGLSDQLAQA 718
```

```
Db 241 DKHTCTPCPAPPELLGGPSVFLFPKPKDTLMISRTPEVTCVVVDVSHEDPEVKFNWYD 300
QY 62 GVEVHNAKTPREEQYNSTYRVVSVLTVLHQDWLNGKEYCKVSNKALPAPIEKTISKAK 121
Db 301 GVEVHNAKTPREEQYNSTYRVVSVLTVLHQDWLNGKEYCKVSNKALPAPIEKTISKAK 360
QY 122 GQPREPQVYTLPPSRDELTKNQVSLTCLVKGFYPSDIAVEWESNGQPENNYKTTTPPVLD 181
Db 361 GQPREPQVYTLPPSRDELTKNQVSLTCLVKGFYPSDIAVEWESNGQPENNYKTTTPPVLD 420
QY 182 DGSFFLYSKLTVDKSRWQGNVFSCVMHEALHNHYTQKSLSLSP- GKDWL 232
Db 421 DGSFFLYSKLTVDKSRWQGNVFSCVMHEALHNHYTQKSLSLSPGSGGLKIAAFNIQT 480
QY 233 -----KAFYDKV-----AEKLKEA 263
Db 661 TFOWLIPDSADTTATPTTHCAIDRIVVAGMLLRGAVVPSALPFOAAYGLSDQLAQA 718

RESULT 7
US-09-784-623-14
; Sequence 14, Application US/09784623
; Patent No. US20020009454A1
; GENERAL INFORMATION:
; APPLICANT: Boone, Thomas C.
; APPLICANT: Hershenson, Susan
; APPLICANT: Bevilacqua, Michael P.
; APPLICANT: Collins, David S.
; TITLE OF INVENTION: COMPOSITION AND METHOD FOR TREATING INFLAMMATORY
; TITLE OF INVENTION: DISEASES
; FILE REFERENCE: A-365F
; CURRENT APPLICATION NUMBER: US/09/784,623
; CURRENT FILING DATE: 2001-02-15
; PRIOR APPLICATION NUMBER: 09/131,247
; PRIOR FILING DATE: 1998-08-07
; PRIOR APPLICATION NUMBER: PCT/US 97/02131
; PRIOR FILING DATE: 1997-02-10
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 14
; LENGTH: 389
; TYPE: PRT
; ORGANISM: Human
US-09-784-623-14

Query Match 86.7%; Score 1240; DB 10; Length 389;
Best Local Similarity 63.7%; Pred. No. 1.6e-31;
Matches 239; Conservative 5; Mismatches 8; Indels 123; Gaps 7;

QY 2 DKHTCTPCPAPPELLGGPSVFLFPKPKDTLMISRTPEVTCVVVDVSHEDPEVKFNWYD 61
Db 10 DKHTCTPCPAPPELLGGPSVFLFPKPKDTLMISRTPEVTCVVVDVSHEDPEVKFNWYD 69
QY 62 GVEVHNAKTPREEQYNSTYRVVSVLTVLHQDWLNGKEYCKVSNKALPAPIEKTISKAK 121
Db 70 GVEVHNAKTPREEQYNSTYRVVSVLTVLHQDWLNGKEYCKVSNKALPAPIEKTISKAK 129
QY 122 GQPREPQVYTLPPSRDELTKNQVSLTCLVKGFYPSDIAVEWESNGQPENNYKTTTPPVLD 181
Db 130 GQPREPQVYTLPPSRDELTKNQVSLTCLVKGFYPSDIAVEWESNGQPENNYKTTTPPVLD 189
```

```
QY 182 DGSFFLYSKLTVDKSRWQGNVFSCVMHEALHNHYTQKSLSLSPCK----- 228
Db 190 DGSFFLYSKLTVDKSRWQGNVFSCVMHEALHNHYTQKSLSLSPCKMRPSGRKSKMQA 249
QY 229 --DW--LKAFY-----DKV----- 238
Db 250 FRWDVNQKTFYLRNNQLVAGYVLOGPNVLEEKIDVVPTEPHALFLGIHGGKMLSCVKS 309
QY 239 -----AEKLKE-----AF-----DW----- 248
Db 310 GDETRIQLEAVNITDLSENKDKKQKRAFIKRDSDGPTTSFSAACPGWFLCTAMEADQPV 369
QY 249 -----LKAFY 253
Db 370 LTNMPDEGVMTKFY 384

RESULT 8
US-09-840-277-2
; Sequence 2, Application US/09840277
; Patent No. US20020158363A1
; GENERAL INFORMATION:
; APPLICANT: FEIGE, ULRICH
; APPLICANT: KOHNO, TADAHIKO
; APPLICANT: LACEY, DAVID LEE
; APPLICANT: BOONE, THOMAS CHARLES
; TITLE OF INVENTION: INTEGRIN/ADHESION ANTAGONISTS
; FILE REFERENCE: A-688A
; CURRENT APPLICATION NUMBER: US/09/840,277
; CURRENT FILING DATE: 2001-08-14
; PRIOR APPLICATION NUMBER: 60/198,919
; PRIOR FILING DATE: 2000-04-21
; PRIOR APPLICATION NUMBER: 60/201,394
; PRIOR FILING DATE: 2000-05-03
; NUMBER OF SEQ ID NOS: 135
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 2
; LENGTH: 228
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-840-277-2

Query Match 86.6%; Score 1238; DB 9; Length 228;
Best Local Similarity 100.0%; Pred. No. 5.2e-32;
Matches 228; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MDKTHCTPCPAPPELLGGPSVFLFPKPKDTLMISRTPEVTCVVVDVSHEDPEVKFNWY 60
Db 1 MDKTHCTPCPAPPELLGGPSVFLFPKPKDTLMISRTPEVTCVVVDVSHEDPEVKFNWY 60
QY 61 DGVEVHNAKTPREEQYNSTYRVVSVLTVLHQDWLNGKEYCKVSNKALPAPIEKTISKA 120
Db 61 DGVEVHNAKTPREEQYNSTYRVVSVLTVLHQDWLNGKEYCKVSNKALPAPIEKTISKA 120
QY 121 KGQPREPQVYTLPPSRDELTKNQVSLTCLVKGFYPSDIAVEWESNGQPENNYKTTTPPVLD 180
Db 121 KGQPREPQVYTLPPSRDELTKNQVSLTCLVKGFYPSDIAVEWESNGQPENNYKTTTPPVLD 180
QY 181 SDGSFFLYSKLTVDKSRWQGNVFSCVMHEALHNHYTQKSLSLSPGK 228
Db 181 SDGSFFLYSKLTVDKSRWQGNVFSCVMHEALHNHYTQKSLSLSPGK 228

RESULT 9
US-09-847-249A-2
; Sequence 2, Application US/09847249A
; Publication No. US20030032588A1
; GENERAL INFORMATION:
; APPLICANT: MARSHALL, WILLIAM S.
; APPLICANT: STARK, KEVIN LEE
; TITLE OF INVENTION: GLUCAGON ANTAGONIST
; FILE REFERENCE: A-693
```

```

: CURRENT APPLICATION NUMBER: US/09/847,249A
:
: CURRENT FILING DATE: 2001-05-02
:
: PRIOR APPLICATION NUMBER: 60/201,436
:
: PRIOR FILING DATE: 2000-05-03
:
: NUMBER OF SEQ ID NOS: 80
:
: SOFTWARE: PatentIn version 3.1
:
: SEQ ID NO 2
:
: LENGTH: 228
:
: TYPE: PRT
:
: ORGANISM: Homo sapiens
:
: US-09-847-249A-2

```

Query Match 86.6%; Score 1238; DB 9; Length 228;  
Best Local Similarity 100.0%; Pred. No. 5.2e-32;  
Matches 228; Conservative 0; Mismatches 0; Indels

QY	1	MDKTHTCPPCPAPBELLGGSVFLFPKPKD	TLMI	STRPEVTCVVVDVSHEDDEVKFNWV	60
Db	1	MDKTHTCPPCPAPBELLGGSVFLFPKPKD	TLMI	STRPEVTCVVVDVSHEDDEVKFNWV	60
QY	61	DGVEVHNAKTKPREEOYNSTRVVS	VLTVLHQDLWLNKGEYKCKVSNKALPAPTEK	TISK	120
Db	61	DGVEVHNAKTKPREEOYNSTRVVS	VLTVLHQDLWLNKGEYKCKVSNKALPAPTEK	TISK	120
QY	121	KGQPREQVYTLPPSRDELTKNOVSL	TVLCKGEGYPSDIAVEWESNQCPENNYK	TTTPPVL	180
Db	121	KGQPREQVYTLPPSRDELTKNOVSL	TVLCKGEGYPSDIAVEWESNQCPENNYK	TTTPPVL	180
QY	181	SDGSFFLYSKLTVDKRWQOQNVFSC	VMHEALHNNHYTKO	KSLSLSPGK	228
Db	181	SDGSFFLYSKLTVDKRWQOQNVFSC	VMHEALHNNHYTKO	KSLSLSPGK	228

```

RESULT 10
US-09-840-669B-2
; Sequence 2, Application US/09840669B
; Publication No. US2003004070A1
; GENERAL INFORMATION:
; APPLICANT: KOHNO, TADAHIKO
; TITLE OF INVENTION: APO-AI/III PEPTIDE DERIVATIVES
; FILE REFERENCE: A-690
; CURRENT APPLICATION NUMBER: US/09/840,669B
; CURRENT FILING DATE: 2002-06-07
; PRIOR APPLICATION NUMBER: 60/198,920
; PRIOR FILING DATE: 2000-04-21
; NUMBER OF SEQ ID NOS: 11
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 2
; LENGTH: 228
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-840-669B-2

```

Query Match 86.6%; Score 1238; DB 9; Length 228;  
Best Local Similarity 100.0%; Pred. No. 5.2e-32;  
Matches 228; Conservative 0; Mismatches 0; Indels

Qy	1	MDKTHTCPPCPAPBELLGGPSVFLFPKPKDFTLMISRTPEVTCVVVDVSHEDPEVKFNWY	60
Db	1	MDKTHTCPPCPAPBELLGGPSVFLFPKPKDFTLMISRTPEVTCVVVDVSHEDPEVKFNWY	60
Qy	61	DGVEVHNNAKTPREEQYNSTIRYVSVLTVLHQDLWLGKEYKCKVSNAKLPAPAEKTI	120
Db	61	DGVEVHNNAKTPREEQYNSTIRYVSVLTVLHQDLWLGKEYKCKVSNAKLPAPAEKTI	120
Qy	121	KGQPREPQVYTLPPSRDELTKNOVSLTCLVKGPYSDIAVEWESNGQPENNYKTTTPVLD	180
Db	121	KGQPREPQVYTLPPSRDELTKNOVSLTCLVKGPYSDIAVEWESNGQPENNYKTTTPVLD	180
Qy	181	SDGSFFLYSKLTVDKSRWQOGNPFSCSVMHAEALHNHYTOKSLSPGK	228
Db	181	SDGSFFLYSKLTVDKSRWQOGNPFSCSVMHAEALHNHYTOKSLSPGK	228

```

RESULT 11
US-09-843-221A-2
; Sequence 2, Application US/09843221A
; Publication No. US20030039654A1
;
; GENERAL INFORMATION:
; APPLICANT: KOSTENIUK, PAUL
; APPLICANT: LIU, CHUAN-FA
; APPLICANT: LACEY, DAVID LEE
;
; TITLE OF INVENTION: MODULATORS OF REIN
; TITLE OF INVENTION: RELATED PROTEIN
; FILE REFERENCE: A-665B
; CURRENT APPLICATION NUMBER: US/09/84
; CURRENT FILING DATE: 2001-04-26
; PRIOR APPLICATION NUMBER: 60/266,673
; PRIOR FILING DATE: 2001-02-06
; PRIOR APPLICATION NUMBER: 60/214,860
; PRIOR FILING DATE: 2000-06-28
; PRIOR APPLICATION NUMBER: 60/200,053
; PRIOR FILING DATE: 2000-04-27
; NUMBER OF SEQ ID NOS: 170
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 2
; LENGTH: 228
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-843-221A-2

```

Query Match 86.6%; Score 1238; DB 9; Length 228;  
Best Local Similarity 100.0%; Pred.No. 5.2e-32;  
Matches 228; Conservative 0; Mismatches 0; Indels

Qy	1	MDKTHTCPPCAPPELLGGPSVFLFPKPKDTLMISRTPEVTCVVVDVSHEDPEVKFNWY	60
Db	1	MDKTHTCPPCAPPELLGGPSVFLFPKPKDTLMISRTPEVTCVVVDVSHEDPEVKFNWY	60
Qy	61	DGSEVHNAKTKPREEQYNSTYRVVSVLTVLRQDWLNGKEYKCKVSKNALPAPIEKTISK	120
Db	61	DGSEVHNAKTKPREEQYNSTYRVVSVLTVLRQDWLNGKEYKCKVSKNALPAPIEKTISK	120
Qy	121	KGQPREPQVYTLPPSRDELTKNOVSLTCLVKGYFSPDIAVEWESNGQPENNYKTTTPVLD	180
Db	121	KGQPREPQVYTLPPSRDELTKNOVSLTCLVKGYFSPDIAVEWESNGQPENNYKTTTPVLD	180
Qy	181	SDGSFFLYSKLTVDKSRWQGNVFSCSVMHEALHNHYTOKSLSPGK	228
Db	181	SDGSFFLYSKLTVDKSRWQGNVFSCSVMHEALHNHYTOKSLSPGK	228

```

RESULT 12
US-09-847-712-2
; Sequence 2, Application US/09847712
; Patent No. US2002009646A1
; GENERAL INFORMATION:
; APPLICANT: LIU, CHUAN-FA
; APPLICANT: MARSHALL, WILLIAM S.
; APPLICANT: REYNOLDS, ANGELA
; TITLE OF INVENTION: CALCIOTONIN-REL
; FILE REFERENCE: A-684
; CURRENT APPLICATION NUMBER: US/09/
; CURRENT FILING DATE: 2001-05-02
; PRIOR APPLICATION NUMBER: 60/201,5
; PRIOR FILING DATE: 2000-05-03
; NUMBER OF SEQ ID NOS: 26
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 2
; LENGTH: 228
; TYPE: prt
; ORGANISM: Homo sapiens
US-09-847-712-2

```

Query Match 86.6%; Score 1238; DB 10; Length 228;  
Best Local Similarity 100.0%; Pred. No. 5.2e-32;

181 SDGSFFLYSKLTVDKSRWQGNVFSCVMHEALHNHYTQKSLSPGK 228

Matches 228; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MDKTHCTPCPCAPPELLGGPSVFLFPKPKDTLMISRTPEVTCVVVDVSHEDPEVKFNWYV 60  
|||||  
Db 1 MDKTHCTPCPCAPPELLGGPSVFLFPKPKDTLMISRTPEVTCVVVDVSHEDPEVKFNWYV 60  
|||||

QY 61 DGVEVHNAKTKPREEQYNSTYRVVSVLTVLHQDLNKGKEYKCKVSNKALPAPIEKTISKA 120  
|||||  
Db 61 DGVEVHNAKTKPREEQYNSTYRVVSVLTVLHQDLNKGKEYKCKVSNKALPAPIEKTISKA 120  
|||||

QY 121 KGQPREPQVYTLPPSRDELTKNOVSLTCLVKGYFSPDSIAVEWESNGQPENNYKTTPPVLD 180  
|||||  
Db 121 KGQPREPQVYTLPPSRDELTKNOVSLTCLVKGYFSPDSIAVEWESNGQPENNYKTTPPVLD 180  
|||||

QY 181 SDGSFFLYSLKLTVDKSRWQOGNVSFCSVMHEALHNHYTOKLSLSPGK 228  
|||||  
Db 181 SDGSFFLYSLKLTVDKSRWQOGNVSFCSVMHEALHNHYTOKLSLSPGK 228  
|||||

RESULT 13  
US-09-840-277-109  
; Sequence 109, Application US/09840277  
; Patent No. US20020168363A1  
; GENERAL INFORMATION:  
; APPLICANT: FEIGE, ULRICH  
; APPLICANT: KOHNO, TADAHIKO  
; APPLICANT: LACEY, DAVID LEE  
; APPLICANT: BOONE, THOMAS CHARLES  
; TITLE OF INVENTION: INTEGRIN/ADHESION ANTAGONISTS  
; FILE REFERENCE: A-688A  
; CURRENT APPLICATION NUMBER: US/09/840,277  
; CURRENT FILING DATE: 2001-08-14  
; PRIOR APPLICATION NUMBER: 60/198,919  
; PRIOR FILING DATE: 2000-04-21  
; PRIOR APPLICATION NUMBER: 60/201,394  
; PRIOR FILING DATE: 2000-05-03  
; NUMBER OF SEQ ID NOS: 135  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 109  
; LENGTH: 282  
; TYPE: PRT  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Echistatin Fc-peptide  
; NAME/KEY: misc.feature  
; LOCATION: (1)..(1)  
; OTHER INFORMATION: Ndel site  
; NAME/KEY: misc.feature  
; LOCATION: (854)..(854)  
; OTHER INFORMATION: BamHI site  
US-09-840-277-109

Query Match 86.6%; Score 1238; DB 9; Length 282;  
Best Local Similarity 100.0%; Pred. No. 8.5e-32;  
Matches 228; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MDKTHCTPCPCAPPELLGGPSVFLFPKPKDTLMISRTPEVTCVVVDVSHEDPEVKFNWYV 60  
|||||  
Db 1 MDKTHCTPCPCAPPELLGGPSVFLFPKPKDTLMISRTPEVTCVVVDVSHEDPEVKFNWYV 60  
|||||

QY 61 DGVEVHNAKTKPREEQYNSTYRVVSVLTVLHQDLNKGKEYKCKVSNKALPAPIEKTISKA 120  
|||||  
Db 61 DGVEVHNAKTKPREEQYNSTYRVVSVLTVLHQDLNKGKEYKCKVSNKALPAPIEKTISKA 120  
|||||

QY 121 KGQPREPQVYTLPPSRDELTKNOVSLTCLVKGYFSPDSIAVEWESNGQPENNYKTTPPVLD 180  
|||||  
Db 121 KGQPREPQVYTLPPSRDELTKNOVSLTCLVKGYFSPDSIAVEWESNGQPENNYKTTPPVLD 180  
|||||

QY 181 SDGSFFLYSLKLTVDKSRWQOGNVSFCSVMHEALHNHYTOKLSLSPGK 228  
|||||  
Db 181 SDGSFFLYSLKLTVDKSRWQOGNVSFCSVMHEALHNHYTOKLSLSPGK 228  
|||||

RESULT 14  
US-09-854-864-31  
; Sequence 31, Application US/09854864  
; Patent No. US20020081296A1  
; GENERAL INFORMATION:  
; APPLICANT: THEILL, LARS EYDE  
; APPLICANT: YU, GANG  
; TITLE OF INVENTION: METHODS AND COMPOSITIONS OF MATTER CONCERNING APRIL/G70, BCWA,  
; TITLE OF INVENTION: BLYS/AGP-3, AND TACI  
; FILE REFERENCE: A-686B  
; CURRENT APPLICATION NUMBER: US/09/854,864  
; CURRENT FILING DATE: 2001-09-11  
; PRIOR APPLICATION NUMBER: US 60/204,039  
; PRIOR FILING DATE: 2000-05-12  
; PRIOR APPLICATION NUMBER: US 60/214,591  
; PRIOR FILING DATE: 2000-06-27  
; NUMBER OF SEQ ID NOS: 31  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 31  
; LENGTH: 394  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-09-854-864-31

Query Match 86.5%; Score 1237; DB 10; Length 394;  
Best Local Similarity 78.6%; Pred. No. 2e-31;  
Matches 231; Conservative 1; Mismatches 5; Indels 57; Gaps 2;

QY 2 DKTHCTPCPCAPPELLGGPSVFLFPKPKDTLMISRTPEVTCVVVDVSHEDPEVKFNWYVD 61  
|||||  
Db 20 DKTHCTPCPCAPPELLGGPSVFLFPKPKDTLMISRTPEVTCVVVDVSHEDPEVKFNWYVD 79  
|||||

QY 62 GVEVHNAKTKPREEQYNSTYRVVSVLTVLHQDLNKGKEYKCKVSNKALPAPIEKTISKAK 121  
|||||  
Db 80 GVEVHNAKTKPREEQYNSTYRVVSVLTVLHQDLNKGKEYKCKVSNKALPAPIEKTISKAK 139  
|||||

QY 122 GQPREPQVYTLPPSRDELTKNOVSLTCLVKGYFSPDSIAVEWESNGQPENNYKTTPPVLD 181  
|||||  
Db 140 GQPREPQVYTLPPSRDELTKNOVSLTCLVKGYFSPDSIAVEWESNGQPENNYKTTPPVLD 199  
|||||

QY 182 DGSFFLYSLKLTVDKSRWQOGNVSFCSVMHEALHNHYTOKLSLSPGKD----- 229  
|||||  
Db 200 DGSFFLYSLKLTVDKSRWQOGNVSFCSVMHEALHNHYTOKLSLSPGKRAVLTKQKKQH 259  
|||||

QY 230 -----WLKA-----FYDKV 238  
| | | | |  
Db 260 SVLHLVPINATSKDDSDVTEVMWQPALRGRGLQAQGYGVRIQDAGVYLLYSQV 313  
| | | | |

RESULT 15  
US-10-215-297-2  
; Sequence 2, Application US/10215297  
; Publication No. US2002019222A1  
; GENERAL INFORMATION:  
; APPLICANT: Blumberg, Richard S.  
; APPLICANT: Simister, Neil E.  
; APPLICANT: Lencer, Wayne I.  
; TITLE OF INVENTION: RECEPTOR SPECIFIC TRANSEPITHELIAL TRANSPORT OF  
; TITLE OF INVENTION: THERAPEUTICS  
; FILE REFERENCE: S1383/7003  
; CURRENT APPLICATION NUMBER: US/10/215,297  
; CURRENT FILING DATE: 2002-08-08  
; PRIOR APPLICATION NUMBER: US 08/578,171  
; PRIOR FILING DATE: 1995-12-29  
; PRIOR APPLICATION NUMBER: US 08/374,159  
; PRIOR FILING DATE: 1995-01-17  
; PRIOR APPLICATION NUMBER: US 09/122,144  
; PRIOR FILING DATE: 1998-07-24  
; NUMBER OF SEQ ID NOS: 22  
; SEQ ID NO 2  
; LENGTH: 22  
; TYPE: PRT  
; ORGANISM: Homo sapiens

US-10-215-297-2

```
Query Match      86.3%; Score 1234; DB 9; Length 229;
Best Local Similarity 99.6%; Pred. No. 7e-32; 0; Indels 0; Gaps 0;
Matches 227; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 MDKTHTCPPCPAPPELLGGPSVFLFPKPKDPTLMISRTPEVTCVVVDVSHEDPEVKFNWYV 60
Db :|||||
QY 2 VDKTHTCPPCPAPPELLGGPSVFLFPKPKDPTLMISRTPEVTCVVVDVSHEDPEVKFNWYV 61
Db :|||||
QY 61 DGEVHNKTKPREEQYNSTYRVSVLTVLHQDLNGKEYCKKVSNAKALPAPIEKTISKA 120
Db :|||||
QY 62 DGEVHNKTKPREEQYNSTYRVSVLTVLHQDLNGKEYCKKVSNAKALPAPIEKTISKA 121
Db :|||||
QY 121 KGQPREPQVYTLPPSRDELTKNOVSLTCLVKGYFPSDIAVEWESNGQPENNYKTTTPVLD 180
Db :|||||
QY 181 SDGSFFLYSKLTVDKSRWQGNVFSCVMHEALHNHYTQKSLSLSPCK 228
Db :|||||
QY 182 SDGSFFLYSKLTVDKSRWQGNVFSCVMHEALHNHYTQKSLSLSPCK 229
Db :|||||
```

Search completed: April 21, 2003, 10:51:21  
Job time : 14.1647 secs



GenCore version 5.1.4.p5\_4578  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: April 21, 2003, 10:42:26 ; Search time 19,4118 Seconds  
(without alignments)  
1307.428 Million cell updates/sec

Title: 2LINK7LINK7  
Perfect score: 1430  
Sequence: 1 MDKTHCTCPCPAPPELLGGPS.....AFDWLKAIFYDKVAEKLKEAF 264

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.0

Searched: 283224 seqs, 96134422 residues  
Total number of hits satisfying chosen parameters: 283224

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000  
Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : PIR\_73:\*  
1: pir1:\*  
2: pir2:\*  
3: pir3:\*  
4: pir4:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	ID	Description
1	1233	86.2	255	4 S31866	Ig gamma-1 chain C
2	1233	86.2	330	1 GHU	Ig gamma-1 chain C
3	1227	85.8	374	2 S69339	Ig heavy chain V r
4	1180	82.5	234	2 P0207	Ig gamma chain C r
5	1158	81.0	377	2 A29511	Ig gamma-3 chain C
6	1156	80.8	377	2 A60764	Ig gamma-3 chain C
7	1146	80.1	326	1 G2HU	Ig gamma-2 chain C
8	1142	79.9	327	1 G4HU	Ig gamma-4 chain C
9	1133	79.2	289	1 G3HUW1	Ig gamma-3 heavy c
10	921	64.4	323	1 GHRB	Ig gamma chain C r
11	917	64.1	328	2 I47160	Ig gamma 2b chain
12	917	64.1	328	2 I47159	Ig gamma 2a chain
13	911	63.7	277	2 I47162	Ig gamma 4 chain c
14	894	62.5	329	1 G2GP	Ig gamma-2 chain C
15	888	62.1	328	2 I47158	Ig gamma 1 chain c
16	885	61.9	328	2 I47161	Ig gamma 3 chain c
17	860	60.1	470	2 S22080	Ig heavy chain pre
18	858	60.0	472	2 S31459	Ig gamma-1 chain -
19	853	59.7	308	2 C30554	Ig heavy chain C r
20	848	59.3	329	1 G3MSC	Ig gamma-3 chain C
21	841	58.8	398	1 G3MSM	Ig gamma-3 chain C
22	840	58.7	333	2 P00018	Ig gamma-2b chain
23	834	58.3	444	2 PC4436	monoclonal antibod
24	827	57.8	393	1 G1MSM	Ig gamma-1 chain C
25	825	57.7	326	2 P00017	Ig gamma-1 chain C
26	823	57.6	324	1 G1MS	Ig gamma-1 chain C
27	819	57.3	399	1 G2MSAM	Ig gamma-1 chain
28	818	57.2	329	2 S00847	Ig gamma-2c chain
29	813	56.9	330	1 G2MSA	Ig gamma-2a chain

30	813	56.9	469	2 S37483	Ig gamma-2a chain
31	809	56.6	335	1 G2MSAB	Ig gamma-2a chain
32	799	55.9	446	2 S40295	Ig gamma-2a chain
33	793	55.5	322	2 PS0019	Ig gamma-2a chain
34	791	55.3	405	1 G2MSBM	Ig gamma-2b chain
35	787	55.0	474	1 G2MS11	Ig gamma-2b chain
36	774	54.1	327	2 S06111	Ig gamma-2 chain C
37	766	53.6	475	2 S01321	Ig gamma-2b chain
38	707	49.4	180	2 I46732	Ig gamma heavy cha
39	602	42.1	218	2 A36040	Ig heavy chain V-I
40	601	42.0	249	2 S69340	Ig heavy chain VHI
41	571	39.9	152	2 S14236	Ig gamma-1 chain C
42	493	34.5	26926	1 I38344	titin, cardiac mus
43	437	30.6	572	2 B45529	Ig y heavy chain (
44	422	29.5	548	2 S38864	Ig epsilon chain C
45	407	28.5	448	2 S03186	Ig heavy chain C r

ALIGNMENTS

RESULT 1  
S31866

Ig gamma-1 chain C region - synthetic  
C:Species: synthetic  
A:Note: Homo sapiens (man) gene engineered and expressed in Escherichia coli  
C:Date: 06-Jan-1995 #sequence\_revision 17-Mar-1997 #text\_change 19-May-2000  
C:Accession: S31866

R:Filpula, D.

submitted to the EMBL Data Library, February 1993

A:Description: Screening method for protein-protein interactions of cloned gene produc

A:Reference number: S31866

A:Accession: S31866

A:Molecule type: mRNA

A:Residues: 1-255 <PIL>

A:Cross-references: EMBL:X70421; NID:g33068; FIDN:CAA49866.1; PID:g33069

C:Keywords: Immunoglobulin

F:1-23/Region: Escherichia coli outer membrane protein A precursor

F:23-255/Region: human Ig gamma-1 chain C region

Query Match 86.2%; Score 1233; DB 4; Length 255;  
Best Local Similarity 100.0%; Pred. No. 2.7e-33;  
Matches 227; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY	2	DKTHTCPCPAPPELLGGPSVFLPFPKPKDTLMISRTPEVTCVVVDVSHEDPEVKFNWYVD	61
DB	29	DKTHTCPCPAPPELLGGPSVFLPFPKPKDTLMISRTPEVTCVVVDVSHEDPEVKFNWYVD	88
QY	62	GVEVHNAKTPREEQNSTYRVVSVLTFLVHODWLNKGKEYKCKVSNKALPAPIEKTISKAK	121
DB	89	GVEVHNAKTPREEQNSTYRVVSVLTFLVHODWLNKGKEYKCKVSNKALPAPIEKTISKAK	148
QY	122	GQPREQVYTLPPSRDELTKNQVSLTCLVKGFYPSDIAVEWESNGQPENNYKTPPPVLD	181
DB	149	GQPREQVYTLPPSRDELTKNQVSLTCLVKGFYPSDIAVEWESNGQPENNYKTPPPVLD	208
QY	182	DGSFFLYSKLTVDKSRWQQGNVSCVMHEALHNNHVTQKSLSPGK	228
DB	209	DGSFFLYSKLTVDKSRWQQGNVSCVMHEALHNNHVTQKSLSPGK	255

RESULT 2  
GHU

Ig gamma-1 chain C region - human

C:Species: Homo sapiens (man)

C:Date: 31-Jan-1981 #sequence\_revision 18-Aug-1982 #text\_change 16-Jul-1999

C:Accession: A93433; S36861; S33887; B90563; A90564; B91668; A91723; A02146

R:Ellison, J.W.; Berson, B.J.; Hood, L.E.

Nucleic Acids Res. 10, 4071-4079, 1982

A:Title: The nucleotide sequence of a human immunoglobulin C-gamma1 gene.

A:Reference number: A93433; MUID:82274238; PMID:6287432

A:Accession: A93433

A:Molecule type: DNA

A:Residues: 1-330 <EL>  
A:Cross-references: EMBL:Z17370  
A:Note: this sequence has the G1m(17) allotypic marker, 97-Igys and the G1m(1) markers,  
A:Note: Lys-330 is removed after translation  
R:Harris, L.J.  
Submitted to the EMBL Data Library, October 1992  
A:Reference number: S33904  
A:Accession: S36861  
A:Molecule type: DNA  
A:Residues: 2-330 <HAR>  
A:Cross-references: EMBL:Z17370  
R:Takahashi, N.; Ueda, S.; Obata, M.; Nikaido, T.; Nakai, S.; Honjo, T.  
Cell 29, 671-679, 1982  
A:Title: Structure of human immunoglobulin gamma genes: Implications for evolution of a  
A:Reference number: S33887; MUID:83001943; PMID:6811139  
A:Accession: S33887  
A:Molecule type: DNA  
A:Residues: 88-113;235-330 <TAK>  
A:Cross-references: EMBL:Z17370  
R:Cunningham, B.A.; Rutishauser, U.; Gall, W.E.; Gottlieb, P.D.; Waxdal, M.J.; Edelman,  
Biochemistry 9, 3161-3170, 1970  
A:Title: The covalent structure of a human gammaG-immunoglobulin. VII. Amino acid sequen  
A:Reference number: A90563; MUID:71064024; PMID:5489771  
A:Contents: myeloma protein Eu  
A:Accession: B90563  
A:Molecule type: protein  
A:Residues: 1-96,'R',98-135 <CUN>  
A:Note: this sequence has the G1m(3) marker, 97-Arg  
R:Rutishauser, U.; Cunningham, B.A.; Bennett, C.; Konigsberg, W.H.; Edelman, G.M.  
Biochemistry 9, 3171-3181, 1970  
A:Title: The covalent structure of a human gammaG-immunoglobulin. VIII. Amino acid sequen  
A:Reference number: A90564; MUID:71064025; PMID:5530842  
A:Contents: Eu  
A:Accession: A90564  
A:Molecule type: protein  
A:Residues: 136-154,'Q',156-165,'Q',167-176,'Q',178-184,'N',196-197,'D',199-238,'E',240,  
A:Note: this sequence has the G1m(non-1) markers, 239-Glu and 241-Met  
R:Ponstingl, H.; Hilschmann, N.  
Hoppe-Seyler's Z. Physiol. Chem. 357, 1571-1604, 1976  
A:Title: Die Primärstruktur eines monoklonalen IgG1-Immunglobulins (Myelomprotein Nie),  
igen Primärstruktur.  
A:Reference number: A91668; MUID:77070269; PMID:826475  
A:Contents: myeloma protein Nie  
A:Accession: B91668  
A:Molecule type: protein  
A:Residues: 1-34,'Q',36-96,'K',98-115,'Q',117-197,'D',199-238,'D',240,'L',242-268,'E',27  
A:Note: this sequence has the G1m(17) and G1m(1) markers  
R:Schmidt, W.E.; Jung, H.D.; Palm, W.; Hilschmann, N.  
Hoppe-Seyler's Z. Physiol. Chem. 364, 713-747, 1983  
A:Title: Die Primärstruktur des kristallisierten monoklonalen Immunglobulins IgG1 KOI  
A:Reference number: A91723; MUID:83289131; PMID:6884994  
A:Contents: myeloma protein KOI; disulfide bonds  
A:Accession: A91723  
A:Molecule type: protein  
A:Residues: 1-96,'R',98-197,'D',199-238,'E',240,'M',242-266,'D',268-271,'D',273-330 <SCH  
A:Note: this sequence has the G1m(3) and G1m(non-1) markers  
R:Gall, W.E.; Edelman, G.M.  
Biochemistry 9, 3188-3196, 1970  
A:Title: The covalent structure of a human gammaG-immunoglobulin. X. Intrachain disulfid  
A:Reference number: A90565; MUID:71064027; PMID:4923144  
A:Contents: annotation; disulfide bonds  
R:Dreker, L.; Schwarz, J.; Reichel, W.; Hilschmann, N.  
Hoppe-Seyler's Z. Physiol. Chem. 357, 1515-1540, 1976  
A:Title: Rule of antibody structure. The primary structure of monoclonal IgG1 immunoglob  
enbromide cleavage products, and the disulfide bridges.  
A:Reference number: A91667; MUID:77070267; PMID:1002129  
A:Contents: annotation; disulfide bonds  
C:Genetics:  
A:Gene: GDB:IGHG1  
A:Cross-references: GDB:120085; OMIM:147100  
A:Map position: 14q32.33-14q32.33  
A:Introns: 99/1; 114/1; 224/1  
C:Complex: An immunoglobulin heterotetramer subunit consists of two identical light (kap

chain disulfide bonds. In some cases, such as IgA and IgM, the subunits associate into  
C:Superfamily: immunoglobulin C region; immunoglobulin homology  
C:Keywords: duplication; glycoprotein; heterotetramer; immunoglobulin  
F:20-85/Domain: immunoglobulin homology <IM1>  
F:137-206/Domain: immunoglobulin homology <IM2>  
F:243-310/Domain: immunoglobulin homology <IM3>  
F:27-83,144-204,250-308/Disulfide bonds: #status experimental  
F:103/Disulfide bonds: interchain (to light chain) #status experimental  
F:109,112/Disulfide bonds: interchain (to heavy chain) #status experimental  
F:180/Binding site: carboxydrate (Asn) (covalent) #status experimental

Query Match 86.2%; Score 1233; DB 1; Length 330;  
Best Local Similarity 100.0%; Pred. No. 6.5e-33;  
Matches 227; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2 DKTHCTPCPAPELLGGPSVFLPFPKPKDLMISRTPEVTCVVVDVSHEDPEVKFNMYVD 61  
|||||  
Db 104 DKTHCTPCPAPELLGGPSVFLPFPKPKDLMISRTPEVTCVVVDVSHEDPEVKFNMYVD 163  
|||||

Qy 62 GVEVHNAKTKPREQYNSTYRVSVLTVTLQDQLNGKREYCKVSNKALPAPIEKTISKAK 121  
|||||  
Db 164 GVEVHNAKTKPREQYNSTYRVSVLTVTLQDQLNGKREYCKVSNKALPAPIEKTISKAK 223  
|||||

Qy 122 GQPREPQVYTLPPSRDELTKNQVSLTCLVKGFYPSDIAVEWESNGQPENNYKTTPPVLD 181  
|||||  
Db 224 GQPREPQVYTLPPSRDELTKNQVSLTCLVKGFYPSDIAVEWESNGQPENNYKTTPPVLD 283  
|||||

Qy 182 DGSFFLSKLTVDKSRWQQGNVSCVMHEALHNNHYTKLSLSPGK 228  
|||||  
Db 284 DGSFFLSKLTVDKSRWQQGNVSCVMHEALHNNHYTKLSLSPGK 330  
|||||

RESULT 3  
S69339  
Ig heavy chain V region precursor - human  
C:Species: Homo sapiens (man)  
C:Date: 19-Mar-1997 #sequence\_revision 19-Mar-1997 #text\_change 01-Dec-2000  
C:Accession: S69339; S72664  
R:Khamilichi, A.A.; Aucouturier, P.; Preud'homme, J.L.; Cogne, M.  
Eur. J. Biochem. 229, 54-60, 1995  
A:Title: Structure of abnormal heavy chains in human heavy-chain-deposition disease.  
A:Reference number: S69339; MUID:95262687; PMID:7744049  
A:Accession: S69339  
A:Status: preliminary  
A:Molecule type: mRNA  
A:Residues: 1-374 <KHA>  
A:Cross-references: EMBL:X81695  
R:Khamilichi, A.A.  
Submitted to the EMBL Data Library, September 1994  
A:Reference number: S72664  
A:Accession: S72664  
A:Status: preliminary  
A:Molecule type: mRNA  
A:Residues: 1-140,'C',142-374 <KHA>  
A:Cross-references: EMBL:X81695  
C:Superfamily: immunoglobulin C region; immunoglobulin homology

Query Match 85.8%; Score 1227; DB 2; Length 374;  
Best Local Similarity 99.1%; Pred. No. 1.6e-32;  
Matches 225; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

Qy 2 DKTHCTPCPAPELLGGPSVFLPFPKPKDLMISRTPEVTCVVVDVSHEDPEVKFNMYVD 61  
|||||  
Db 148 DKTHCTPCPAPELLGGPSVFLPFPKPKDLMISRTPEVTCVVVDVSHEDPEVKFNMYVD 207  
|||||

Qy 62 GVEVHNAKTKPREQYNSTYRVSVLTVTLQDQLNGKREYCKVSNKALPAPIEKTISKAK 121  
|||||  
Db 208 GVEVHNAKTKPREQYNSTYRVSVLTVTLQDQLNGKREYCKVSNKALPAPIEKTISKAK 267  
|||||

Qy 122 GQPREPQVYTLPPSRDELTKNQVSLTCLVKGFYPSDIAVEWESNGQPENNYKTTPPVLD 181  
|||||  
Db 268 GQPREPQVYTLPPSRDELTKNQVSLTCLVKGFYPSDIAVEWESNGQPENNYKTTPPVLD 327  
|||||



QY 182 DGSFFLYSKLTVDKSRWQGNVFSVMSVHEALHNHYTKSLSPGK 228  
 |||||  
 Db 328 DGSFFLYSKLTVDKSRWQGNVFSVMSVHEALHNHYTKSLSPGK 374

## RESULT 4

PT0207  
 Ig gamma chain C region - chimpanzee  
 C:Species: Pan troglodytes (chimpanzee)  
 C:Date: 23-Nov-1991 #sequence\_revision 23-Nov-1991 #text\_change 16-Jul-1999  
 C:Accession: PT0207  
 R:Ellrich, P.H.; Moustafa, Z.A.; Oestberg, L.  
 Mol. Immunol. 28, 319-322, 1991  
 A:Title: Nucleotide sequence of chimpanzee Fc and hinge regions.  
 A:Reference number: PT0207; MUID:91287716; PMID:2062315  
 A:Accession: PT0207  
 A:Molecule type: mRNA  
 A:Residues: 1-234 <EHR>  
 C:Superfamily: immunoglobulin C region; immunoglobulin homology  
 C:Keywords: immunoglobulin  
 F:48-117/Domain: immunoglobulin homology <IMM>

Query Match 82.5%; Score 1180; DB 2; Length 234;  
 Best Local Similarity 98.8%; Pred. No. 1.1e-31;  
 Matches 217; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 2 DKTHTCPPCAPPELLGSPSVFLPPKPKDTLMISRTPEVTCVVVDVSHEDPEVKFNWYVD 61  
 |||||

Db 15 DTHTCPPCAPPELLGSPSVFLPPKPKDTLMISRTPEVTCVVVDVSHEDPEVKFNWYVD 74  
 |||||

QY 62 GVEVHNAKTRPREQYNSTYRVVSVLTVLHQDWLNGKEYKCKVSNKALPAPIEKTISKAK 121  
 |||||

Db 75 GVEVHNAKTRPREQYNSTYRVVSVLTVLHQDWLNGKEYKCKVSNKALPAPIEKTISKAK 134  
 |||||

QY 122 GQREPOVYTLPPSRDELTKNQVSLTCLVKGFYPSDIAVEWESNGQPENNYKTTTPVLD 181  
 |||||

Db 135 GQREPOVYTLPPSRDELTKNQVSLTCLVKGFYPSDIAVEWESNGQPENNYKTTTPVLD 194  
 |||||

QY 182 DGSFFLYSKLTVDKSRWQGNVFSVMSVHEALHNHYTKS 221  
 |||||

Db 195 DGSFFLYSKLTVDKSRWQGNVFSVMSVHEALHNHYTKS 234  
 |||||

## RESULT 5

A23511  
 Ig gamma-3 chain C region (allotype G3m(b)) - human  
 C:Species: Homo sapiens (man)  
 C:Date: 28-Dec-1987 #sequence\_revision 28-Dec-1987 #text\_change 23-Jul-1999  
 C:Accession: A23511  
 R:Huck, S.; Fort, P.; Crawford, D.H.; Lefranc, M.P.; Lefranc, G.  
 Nucleic Acids Res. 14, 1779-1789, 1986  
 A:Title: Sequence of a human immunoglobulin gamma 3 heavy chain constant region gene: cDNA  
 A:Reference number: A23511; MUID:86148507; PMID:3081877  
 A:Accession: A23511  
 A:Molecule type: DNA  
 A:Residues: 1-377 <HUC>  
 A:Cross-references: GB:X03604; GB:M12958; NID:g33070; PIDN:CAA27268.1; PID:g577056  
 C:Genetics:  
 A:Gene: GDB:IGHG3  
 A:Cross-references: GDB:119339; OMIM:147120  
 A:Map position: 14q32.33-14q32.33  
 A:Introns: 98/3; 115/3; 130/3; 145/3; 160/3; 270/3  
 C:Superfamily: immunoglobulin C region; immunoglobulin homology  
 C:Keywords: immunoglobulin  
 F:20-85/Domain: immunoglobulin homology <IMM>

Query Match 81.0%; Score 1158; DB 2; Length 377;  
 Best Local Similarity 78.3%; Pred. No. 3e-30;  
 Matches 213; Conservative 8; Mismatches 6; Indels 45; Gaps 1;

QY 2 DKTHT-----  
 |||||  
 Db 106 DTHTCPRCPKSCDTPPPCPRCPKSCDTPPPCPRCPKSCDTPPPCPRCPAPPELL 165

QY 17 GGPSVFLFPKPKDTLMISRTPEVTCVVVDVSHEDPEVKFNWYVDGVEVHNAKTKPREEQ 76  
 |||||  
 Db 166 GGPSVFLFPKPKDTLMISRTPEVTCVVVDVSHEDPEVKFNWYVDGVEVHNAKTKPREEQ 225  
 |||||  
 QY 77 YNSTYRVVSVLTVLHQDWLNGKEYKCKVSNKALPAPIEKTISKAKGQPREPQVYTLPPSR 136  
 |||||  
 Db 226 YNSTFRVSVLTVLHQDWLNGKEYKCKVSNKALPAPIEKTISKAKGQPREPQVYTLPPSR 285  
 |||||  
 QY 137 DELTKNQVSLTCLVKGFYPSDIAVEWESNGQPENNYKTTTPVLDSDGSGFFLYSKLTVDKS 196  
 |||||  
 Db 286 EEMTKNQVSLTCLVKGFYPSDIAVEWESNGQPENNYKTTTPVLDSDGSGFFLYSKLTVDKS 345  
 |||||  
 QY 197 RWQGNVFSVMSVHEALHNHYTKSLSPGK 228  
 |||||  
 Db 346 RWQGNVFSVMSVHEALHNHYTKSLSPGK 377  
 |||||

## RESULT 6

A60764  
 Ig gamma-3 chain C region, form LAT - human  
 C:Species: Homo sapiens (man)  
 C:Date: 14-May-1993 #sequence\_revision 14-May-1993 #text\_change 16-Jul-1999  
 C:Accession: A60764  
 R:Huck, S.; Lefranc, G.; Lefranc, M.P.  
 Immunogenetics 30, 250-257, 1989  
 A:Title: A human immunoglobulin IGHG3 allele (Gmb0, bl, c3, c5, u) with an IGHG4 con-

A:Reference number: A60764; MUID:90007613; PMID:2571587  
 A:Accession: A60764  
 A:Status: preliminary  
 A:Molecule type: DNA  
 A:Residues: 1-377 <HUC>  
 C:Superfamily: immunoglobulin C region; immunoglobulin homology  
 C:Keywords: immunoglobulin  
 F:20-85/Domain: immunoglobulin homology <IMM>

Query Match 80.8%; Score 1156; DB 2; Length 377;  
 Best Local Similarity 78.3%; Pred. No. 3.5e-30;  
 Matches 213; Conservative 8; Mismatches 6; Indels 45; Gaps 1;

QY 2 DKTHT-----  
 |||||  
 Db 106 DTHTCPRCPKSCDTPPPCPRCPKSCDTPPPCPRCPKSCDTPPPCPRCPAPPELL 165  
 |||||

QY 17 GGPSVFLFPKPKDTLMISRTPEVTCVVVDVSHEDPEVKFNWYVDGVEVHNAKTKPREEQ 76  
 |||||

Db 166 GGPSVFLFPKPKDTLMISRTPEVTCVVVDVSHEDPEVKFNWYVDGVEVHNAKTKPREEQ 225  
 |||||

QY 77 YNSTYRVVSVLTVLHQDWLNGKEYKCKVSNKALPAPIEKTISKAKGQPREPQVYTLPPSR 136  
 |||||

Db 226 YNSTFRVSVLTVLHQDWLNGKEYKCKVSNKALPAPIEKTISKAKGQPREPQVYTLPPSR 285  
 |||||

QY 137 DELTKNQVSLTCLVKGFYPSDIAVEWESNGQPENNYKTTTPVLDSDGSGFFLYSKLTVDKS 196  
 |||||

Db 286 EEMTKNQVSLTCLVKGFYPSDIAVEWESNGQPENNYKTTTPVLDSDGSGFFLYSKLTVDKS 345  
 |||||

QY 197 RWQGNVFSVMSVHEALHNHYTKSLSPGK 228  
 |||||

Db 346 RWQGNVFSVMSVHEALHNHYTKSLSPGK 377  
 |||||

## RESULT 7

G2HU  
 Ig gamma-2 chain C region - human  
 C:Species: Homo sapiens (man)  
 C:Date: 30-Apr-1981 #sequence\_revision 13-Jun-1983 #text\_change 21-Jul-2000  
 C:Accession: A93906; A92809; A90752; A93132; A02148  
 R:Ellison, J.; Hood, L.  
 Proc. Natl. Acad. Sci. U.S.A. 79, 1984-1988, 1982  
 A:Title: Linkage and sequence homology of two human immunoglobulin gamma heavy chain  
 A:Reference number: A93906; MUID:82197621; PMID:6804948  
 A:Accession: A93906  
 A:Molecule type: DNA

A:Residues: 1-326 <ELL>  
A:Cross-references: GB:V00554; GB:J00230; NID:g32759; PID: CAB58438.1; PID:g606056  
A:Note: Lys-326 is probably removed posttranslationally  
R:Wang, A.C.; Tung, E.; Fudenberg, H.H.  
J. Immunol. 125, 1048-1054, 1980  
A:Title: The primary structure of a human IgG2 heavy chain: genetic, evolutionary, and  
A:Reference number: A92809; MUID:81007873; PMID:6774012  
A:Contents: myeloma protein til  
A:Accession: A92809  
A:Molecule type: protein  
A:Residues: 1-19, 'Q', 21-57, 'Z', 59, 'A', 61-193, 'D', 195-325 <NAN>  
A:Note: Trp-156 is at or near the complement-binding site  
R:Connell, G.E.; Parr, D.M.; Hofmann, T.  
Can. J. Biochem. 57, 758-767, 1979  
A:Title: The amino acid sequences of the three heavy chain constant region domains of a  
A:Reference number: A90752; MUID:80001357; PMID:113060  
A:Contents: myeloma protein zle  
A:Accession: A90752  
A:Molecule type: protein  
A:Residues: 1-24, 'E', 26-57, 'EV', 60-85, 132-171, 'ZZZ', 175, 'B', 177-193, 'D', 195-196, 'Q', 198-  
A:Note: This sequence has since been revised  
R:Hofmann, T.; Parr, D.M.  
Mol. Immunol. 16, 923-925, 1979  
A:Title: A note on the amino acid sequence of residues 381-391 of human immunoglobulin G  
A:Reference number: A93132; MUID:80114419; PMID:118920  
A:Contents: zle  
A:Accession: A93132  
A:Molecule type: protein  
A:Residues: 238-275 <HOF>  
R:Hofmann, T.; Parr, D.M.  
submitted to the Atlas, March 1980  
A:Reference number: A94591  
A:Contents: annotation; zle, revisions to residues 25, 59, 60, and 264-268  
A:Note: the revised sequence differs from that shown in having 60-Ala and in the amidati  
ned  
R:Milstein, C.; Frangione, B.  
Biochem. J. 121, 217-225, 1971  
A:Title: Disulfide bridges of the heavy chain of human immunoglobulin G2.  
A:Reference number: A90253; MUID:72033500; PMID:4940472  
A:Contents: annotation; myeloma protein Sa, disulfide bonds  
R:Frangione, B.; Milstein, C.; Pink, J.R.L.  
Nature 221, 145-148, 1969  
A:Title: Structural studies of immunoglobulin G.  
A:Reference number: A93157; MUID:69064124; PMID:5782707  
A:Contents: annotation; Sa, disulfide bonds  
C:Genetics:  
A:Gene: GDB:IGHG2  
A:Cross-references: GDB:119338; OMIM:147110  
A:Map position: 14q32.33-14q32.33  
C:Complex: An immunoglobulin heterotetramer subunit consists of two identical light (kap  
hain disulfide bonds. In some cases, such as IgA and IgM, the subunits associate into la  
C:Superfamily: immunoglobulin C region; immunoglobulin homology  
C:Keywords: duplication; glycoprotein; heterotetramer; immunoglobulin  
F:20-85/Domain: immunoglobulin homology <IM1>  
F:133-202/Domain: immunoglobulin homology <IM2>  
F:239-306/Domain: immunoglobulin homology <IM3>  
F:14/Disulfide bonds: interchain (to light chain) #status experimental  
F:27-83,141-201,246-304/Disulfide bonds: #status experimental  
F:102,103,106,109/Disulfide bonds: interchain (to heavy chain) #status experimental  
F:176/Binding site: carbohydrate (Asn) (covalent) #status predicted  
Query Match 80.1%; Score 1146; DB 1; Length 326;  
Best Local Similarity 90.6%; Pred. No. 4.6e-30;  
Matches 212; Conservative 9; Mismatches 6; Indels 7; Gaps 2;  
QY 1 MDKT-----HPCPCPAPELLGGPSVFLFPPKPKDTLMISRTPEVTCVVDVDSHEDPEV 54  
DB 94 VDKTVERCCVCCPCPCAPP-VAGSVFLFPPKPKDTLMISRTPEVTCVVDVDSHEDPEV 152  
QY 55 KFNWYVDGVGVHNAKTPREEQYNSTYRVSVLTVLDHQLNGKEVKCKVSNKALPAPIE 114  
DB 153 QFNWYVDGVGVHNAKTPREEQNFSTYRVSVLTVLDHQLNGKEVKCKVSNKGLPAPIE 212

QY 115 KTISKAKGPREPQVYTLPPSRDELTKNOVSLTCLVKGFYPSDIAVEWESNGQPENNYKT 174  
DB 213 KTISKAKGPREPQVYTLPPSRDEMTKNQVSLTCLVKGFYPSDIAVEWESNGQPENNYKT 272  
QY 175 TTPVLDSGDSFFLYSKLTVDKSRWQQGNVSCSMHEALHNHYTQKSLSLSPGK 228  
DB 273 TTPMLDSGDSFFLYSKLTVDKSRWQQGNVSCSMHEALHNHYTQKSLSLSPGK 326  
RESULT 8  
GAMU  
Ig gamma-4 chain C region - human  
C:Species: Homo sapiens (man)  
C:Date: 02-Apr-1982 #sequence\_revision 02-Apr-1982 #text\_change 16-Jul-1999  
C:Accession: A90933; A90249; A02150  
R:Ellison, J.; Buxbaum, J.; Hood, L.  
DNA 1, 11-18, 1981  
A:Title: Nucleotide sequence of a human immunoglobulin C-gamma4 gene.  
A:Reference number: A90933; MUID:83157104; PMID:6299662  
A:Accession: A90933  
A:Molecule type: DNA  
A:Residues: 1-327 <ELL>  
A:Note: The sequence was determined from the germline gene  
R:Pink, J.R.L.; Buttery, S.H.; De Vries, G.M.; Milstein, C.  
Biochem. J. 117, 33-47, 1970  
A:Title: Human immunoglobulin subclasses. Partial amino acid sequence of the constan  
A:Reference number: A90249; MUID:70207560; PMID:4192699  
A:Accession: A90249  
A:Molecule type: protein  
A:Residues: 1-30; 81-326 <PIN>  
C:Genetics:  
A:Gene: GDB:IGHG4  
A:Cross-references: GDB:119340; OMIM:147130  
A:Map position: 14q32.33-14q32.33  
A:Introns: 99/1; 111/1; 221/1  
C:Complex: An immunoglobulin heterotetramer subunit consists of two identical light (l  
hain disulfide bonds. In some cases, such as IgA and IgM, the subunits associate into  
C:Superfamily: immunoglobulin C region; immunoglobulin homology  
C:Keywords: duplication; glycoprotein; heterotetramer; immunoglobulin  
F:20-85/Domain: immunoglobulin homology <IM1>  
F:99-110/Region: hinge  
F:134-203/Domain: immunoglobulin homology <IM2>  
F:240-307/Domain: immunoglobulin homology <IM3>  
F:14/Disulfide bonds: interchain (to light chain) #status experimental  
F:27-83,141-201,247-305/Disulfide bonds: #status predicted  
F:106,109/Disulfide bonds: interchain (to heavy chain) #status experimental  
F:177/Binding site: carbohydrate (Asn) (covalent) #status predicted  
Query Match 79.9%; Score 1142; DB 1; Length 327;  
Best Local Similarity 84.7%; Pred. No. 6.3e-30;  
Matches 211; Conservative 9; Mismatches 6; Indels 23; Gaps 1;  
QY 3 KTHT-----CPCPAPELLGGPSVFLFPPKPKDTLMISRTPE 39  
DB 79 KTYCNVDHKNPKNTKVKRVESKYGPPCPSPAPEFLGGPSVFLFPPKPKDTLMISRTPE 138  
QY 40 VTCVVVDVSHEDPEVKFNWYVDGVEVHNAKTPREEQYNSTYRVSVLTVLDHQLNGKE 99  
DB 139 VTCVVVDVSHEDPEVKFNWYVDGVEVHNAKTPREEQYNSTYRVSVLTVLDHQLNGKE 198  
QY 100 YKCKVSNKALPAPIEKTISKAKGQPREPOVYTLPPSRDELTKNQVSLTCLVKGFYPSDIA 159  
DB 199 YKCKVSNKGLPSSIEKTISKAKGQPREPOVYTLPPSRDELTKNQVSLTCLVKGFYPSDIA 258  
QY 160 VEWESNGQPENNYKTTPPVLDSDGSFFLYSKLTVDKSRWQQGNVSCSMHEALHNHYTQ 219  
DB 259 VEWESNGQPENNYKTTPPVLDSDGSFFLYSKLTVDKSRWQQGNVSCSMHEALHNHYTQ 318  
QY 220 KSLSLSPGK 228  
DB 319 KSLSLSLGK 327

## RESULT 9

G3HUWI  
Ig gamma-3 heavy chain disease proteins - human  
C:Species: Homo sapiens (man)  
C:Date: 31-Dec-1979 #sequence\_revision 23-Oct-1981 #text\_change 16-Jul-1999  
C:Accession: A90442; A92219; A90198; A93915; A02149  
R:Frangione, B.; Rosenwasser, E.; Prelli, F.; Franklin, E.C.  
Biochemistry 19, 4304-4308, 1980  
A:Title: Primary structure of human gamma3 immunoglobulin deletion mutant: gamma3 heavy-  
A:Reference number: A90442; MUID:81021548; PMID:6774747  
A:Contents: heavy chain disease protein wis  
A:Accession: A90442  
A:Molecule type: protein  
A:Residues: 1-289 <FRA>  
A:Note: the molecule is a dimer linked by 12 disulfide bonds; it has an extra interchain  
A:Note: this protein lacks most of the V region and all of the CH1 region. Residue 12 cd  
A:Note: the sequence of residues 42-76 was taken from the reference that follows  
R:Michaelson, T.E.; Frangione, B.; Franklin, E.C.  
J. Biol. Chem. 252, 883-889, 1977  
A:Title: Primary structure of the 'hinge' region of human IgG3. Probable quadruplication  
A:Reference number: A92219; MUID:77118561; PMID:402363  
A:Contents: normal gamma-3 chains, sequence corresponding to residues 12-97 of protein w  
A:Accession: A92219  
A:Molecule type: protein  
A:Residues: 12-97 <MIC>  
A:Note: the hinge region in gamma-3 chains is about four times as long as in other gamma  
idue segment (12-28)  
A:Note: cysteines at positions 24, 27, 33, 39, 42, 48, 54, 57, 63, 69, and 72 form inter  
R:Wolfenstein-Todel, C.; Frangione, B.; Prelli, F.; Franklin, E.C.  
Biochem. Biophys. Res. Commun. 71, 907-914, 1976  
A:Title: The amino acid sequence of "heavy chain disease" protein ZUC. Structure of the  
A:Reference number: A90198; MUID:77021516; PMID:823945  
A:Contents: heavy chain disease protein Zuc, partial sequence corresponding to residues  
A:Accession: A90198  
A:Molecule type: protein  
A:Residues: 59-125, 'EB', 128-226, 228-289 <WOL>  
A:Note: this protein lacks most of the V region, all of the CH1 region, and part of the  
R:Alexander, A.; Steinmetz, M.; Barritault, D.; Frangione, B.; Franklin, E.C.; Hood, L.;  
Proc. Natl. Acad. Sci. U.S.A. 79, 3260-3264, 1982  
A:Title: gamma heavy chain disease in man: cDNA sequence supports partial gene deletion  
A:Reference number: A93915; MUID:82247835; PMID:6808505  
A:Contents: heavy chain disease protein Omn  
A:Accession: A93915  
A:Molecule type: mRNA  
A:Residues: 12-70; 72-114; 116-125, 'E', 127-133, 'L', 135-136, 'E', 138, 'Y', 140-154, 'D', 156-157  
A:Note: a carboxyl-terminal Lys is removed posttranslationally  
A:Note: this sequence may represent an allelic form or another gamma chain subclass  
C:Comment: The heavy chain disease protein wis is shown.  
C:Genetics:  
A:Gene: GDB:IGHG3  
A:Cross-references: GDB:119339; OMIM:147120  
A:Map position: 14q32.33-14q32.33  
C:Superfamily: immunoglobulin C region; immunoglobulin homology  
C:Keywords: duplication; glycoprotein; immunoglobulin; pyroglutamic acid  
F:203-270/Domain: immunoglobulin homology <IMM>  
F:1/Modified site: pyrrolidone carboxylic acid (Gln) #status experimental  
F:6,140/Binding site: carbohydrate (Asn) (covalent) #status experimental  
Query Match 79.2%; Score 1133; DB 1; Length 289;  
Best Local Similarity 76.4%; Pred. No. 8.3e-30;  
Matches 207; Conservative 13; Mismatches 6; Indels 45; Gaps 1;  
QY 2 DKTH-----CPCPAPELL 16  
Db 19 DTHTCRCPKSCDTPPCRCRCPKSCDTPPCRCRCPKSCDTPPCRCPAPELL 78  
QY 17 GGPVFLFPKPKDLMISRTPEVTCVVVDVSHEDPEVKFNKYVDGVGHVHNAKTPREEQ 76  
Db 79 GGPVFLFPKPKDLMISRTPEVTCVVVDVSHEDPEVKFYVDGVGHVHNAKTPREEQ 138  
QY 77 YNSTYRVSVLTUHQDWLNGKEYCKVKSNKALPAPIETKISKAKGQRPQVYTLPPSR 136  
Db 139 FNSTFRVSVLTUHQDWLNGKEYCKVKSNKALPAPIETKISKAKGQRPQVYTLPPSR 198

QY 137 DELTKNOVSLTCLVKGYPSDIAVWESNGOPENNYKTTTPVLDSDGSFFLYSKLTVDKS 196  
Db 199 EEMTKNOVSLTCLVKGYPSDIAVWESNGOPENNYKTTTPVLDSDGSFFLYSKLTVDKS 258  
QY 197 RWQOGNFCVSMHEALHNHYTKLSLSPG 227  
Db 259 RWQOGNFCVSMHEALHNHYTKLSLSPG 289

## RESULT 10

## GHRB

Ig gamma chain C region - rabbit  
C:Species: Oryctolagus cuniculus (domestic rabbit)  
C:Date: 24-Apr-1984 #sequence\_revision 15-Nov-1984 #text\_change 16-Jul-1999  
C:Accession: A91749; A90290; A93928; A90245; A94416; A02161  
R:Bernstein, K.E.; Alexander, C.B.; Mage, R.G.  
Immunogenetics 18, 387-397, 1983  
A:Title: Nucleotide sequence of a rabbit IgG heavy chain from the recombinant F-I ha  
A:Reference number: A91749; MUID:84030930; PMID:6313520  
A:Accession: A91749  
A:Molecule type: mRNA  
A:Residues: 1-323 <BER>  
A:Note: this sequence has the d12 allotypic marker, 104-Thr, and the e14 marker, 185  
R:Pratt, D.M.; Mole, L.E.  
Biochem. J. 151, 337-349, 1975  
A:Title: Sequence studies on the constant region of the Fd sections of rabbit immuno  
A:Reference number: A90290; MUID:76135469; PMID:1243651  
A:Accession: A90290  
A:Molecule type: protein  
A:Residues: 1-47, 'E', 49-71, 'PV', 72-128 <PRA>  
R:Martens, C.L.; Moore, K.W.; Steinmetz, M.; Hood, L.; Knight, K.L.  
Proc. Natl. Acad. Sci. U.S.A. 79, 6018-6022, 1982  
A:Title: Heavy chain genes of rabbit IgG; isolation of a cDNA encoding gamma heavy c  
A:Reference number: A93928; MUID:83299917; PMID:6193512  
A:Accession: A93928  
A:Molecule type: mRNA  
A:Residues: 88-103, 'M', 105-143, 'E', 145-184, 'A', 186, 'E', 188-266 <MAR>  
A:Cross-references: GB:M16426; NID:G165111; PID:AAA31289.1; PID:G165112  
A:Note: this sequence has the d11 allotypic marker, 104-Met, and the e15 allotypic m  
R:Fruchter, R.G.; Jackson, S.A.; Mole, L.E.; Porter, R.R.  
Biochem. J. 116, 249-259, 1970  
A:Title: Sequence studies of the Fd section of the heavy chain of rabbit immunoglob  
A:Reference number: A90245; MUID:70110015; PMID:5461106  
A:Accession: A90245  
A:Molecule type: protein  
A:Residues: 132-143, 'E', 145-161 <PRU>  
R:Hill, R.L.; Lebovitz, H.E.; Fellows Jr., R.E.; Delaney, R.  
in Gamma Globulins, Nobel Symp. 3, Killander, J., ed., pp.109-127, Almqvist and Wiks  
A:Reference number: A94416  
A:Accession: A94416  
A:Molecule type: protein  
A:Residues: 129-131; 155-172, 'D', 174-184, 'A', 186, 'E', 188-200, 'D', 202-217, 'E', 219-232,  
A:Note: this has the e15 allotypic marker, 185-Ala  
C:Complex: An immunoglobulin heterotrimer subunit consists of two identical light  
hain disulfide bonds. In some cases, such as IgA and IgM, the subunits associate int  
C:Superfamily: immunoglobulin C region; immunoglobulin homology  
C:Keywords: duplication; glycoprotein; heterotrimer; immunoglobulin  
F:20-82/Domain: immunoglobulin homology <IMI>  
F:130-199/Domain: immunoglobulin homology <IM2>  
F:236-303/Domain: immunoglobulin homology <IM3>  
F:173/Binding site: carbohydrate (Asn) (covalent) #status predicted  
Query Match 54.4%; Score 921; DB 1; Length 323;  
Best Local Similarity 71.7%; Pred. No. 1.2e-22;  
Matches 167; Conservative 29; Mismatches 32; Indels 5; Gaps 2;  
QY 1 MDXT---HTC--PPCAPELLGGPSVLFPPKPKDTLMISRTPEVTCVVVDVSHEDPEVK 55  
Db 91 VDKTVASTCKPTCPPELLGGPSVLFPPKPKDTLMISRTPEVTCVVVDVSHEDPEVK 150  
QY 56 FNMVYVDGVEVHNAKTPREEQNSTYRVSVYLTUHQDWLNGKEYCKVKSNKALPAPIEK 115

Db 151 FTWYINNEQVTRAPPLREQOQNSTIRVWSTLPIHQDWLRGKFKCKVHNKALPAPIEK 210  
QY 116 TISKAKGQPREPOVYTLPPSRDELTKNOVSLTCLVKGFPDIAVWESNGQPPENNYKTT 175  
Db 211 TISKARQOPLEPKVYVYNGPRELSRSVSLTCLVKGFPDIAVWESNGKRAEDNYKT 270  
QY 176 PVLDSGDFLYSKLTVDKSRQOQNVFSCVMHEALHNHYTKSLSPGK 228  
Db 271 PAVLDSGDFLYSKLTVDKSRQOQNVFSCVMHEALHNHYTKSLSPGK 323

RESULT 11  
I47160  
Ig gamma 2b chain constant region - pig (fragment)  
C:Species: Sus scrofa domestica (domestic pig)  
C:Date: 21-Feb-1997 #sequence\_revision 21-Feb-1997 #text\_change 21-Jan-2000  
C:Accession: I47160  
R:Kacskovics, I.; Sun, J.; Butler, J.E.  
J. Immunol. 153, 3565-3573, 1994  
A:Title: Five putative subclasses of swine IgG identified from the cDNA sequences of a  
A:Reference number: I47158; MUID:95015845; PMID:7930579  
A:Accession: I47160  
A:Status: preliminary; translated from GB/EMBL/DBDJ  
A:Molecule type: mRNA  
A:Residues: 1-328 <KAC>  
A:Cross-references: EMBL:U03780; NID:g433125; PIDN:AAAS2218.1; PID:g433126  
C:Genetics:  
A:Gene: IgG2b  
C:Superfamily: immunoglobulin C region; immunoglobulin homology  
F:133-202/Domain: immunoglobulin homology <IMM>

Query Match 64.1%; Score 917; DB 2; Length 328;  
Best Local Similarity 65.4%; Pred. No. 1.7e-22;  
Matches 166; Conservative 32; Mismatches 24; Indels 32; Gaps 4;

QY 3 KTHTC-----PPCP-----APELLGSPVFLFPKPKDTLMISR 134  
Db 79 KSYTCNVNHPATTTKVDKRVGTRKPPCPICPACESP-----GPSVFIFPPKPKDTLMISR 134

QY 37 TPEVTCVVVDVSHEDPEVKFNMYVDGVEVHNAKTKPREEQNSTYRVVSVLTVLHQDWL 96  
Db 135 TPQVTCVVVDVSENVDFVFNVDGVEVHTAQTREKEQFNSTYRVVSVLTVLHQDWL 194

QY 97 GREYCKVSNKALPAPIETISKAKGQPREPOVYTLPPSRDELTKNOVSLTCLVKGFP 156  
Db 195 GKEFKCKVNNKDLPAITRIISKAKGQTPREPOVYTLPPHAELSRKSVITCLVIGF 254

QY 157 DIAVESNGQ--PENNYKTTTPVLDSDGFFLYSKLTVDKSRQOQNVFSCVMHEALH 214  
Db 255 DIDVEQRNGQPEPEGNRYRTTPQDDVDGTYFLYSKFSVDKASWQGGGIFQCAVMHEALH 314

QY 215 NYHTOKSLSPGK 228  
Db 315 NYHTOKSLSPGK 328

RESULT 13  
I47162  
Ig gamma 4 chain constant region - pig (fragment)  
C:Species: Sus scrofa domestica (domestic pig)  
C:Date: 21-Feb-1997 #sequence\_revision 21-Feb-1997 #text\_change 21-Jan-2000  
C:Accession: I47162  
R:Kacskovics, I.; Sun, J.; Butler, J.E.  
J. Immunol. 153, 3565-3573, 1994  
A:Title: Five putative subclasses of swine IgG identified from the cDNA sequences of  
A:Reference number: I47158; MUID:95015845; PMID:7930579  
A:Accession: I47162  
A:Status: preliminary; translated from GB/EMBL/DBDJ  
A:Molecule type: mRNA  
A:Residues: 1-277 <KAC>  
A:Cross-references: EMBL:U03782; NID:g433129; PIDN:AAAS2220.1; PID:g433130  
C:Genetics:  
A:Gene: IgG4  
C:Superfamily: immunoglobulin C region; immunoglobulin homology  
F:82-151/Domain: immunoglobulin homology <IMM>

Query Match 63.7%; Score 911; DB 2; Length 277;  
Best Local Similarity 66.4%; Pred. No. 1.5e-22;  
Matches 166; Conservative 31; Mismatches 29; Indels 24; Gaps 4;

QY 3 KTHTC-----PPCP-----APELLG-SPVFLFPKPKDTLMISRTPEV 40  
Db 28 KSYTCNVNHPATTTKVDKRVGTRKPPCPICPACESP-----GPSVFIFPPKPKDTLMISRTPKV 87

QY 41 TCVVVDVSHEDPEVKFNMYVDGVEVHNAKTKPREEQNSTYRVVSVLTVLHQDWLNGKEY 100  
Db 88 TCVVVDVSENVDFVFNVDGVEVHTAQTREKEQFNSTYRVVSVLTVLHQDWLNGKEY 147

QY 101 KCKVSNKALPAPIETISKAKGQPREPOVYTLPPSRDELTKNOVSLTCLVKGFPDIAV 160  
Db 148 KCKVNNKDLPAITRIISKAKGQTPREPOVYTLPPTEELSRKSVITCLVIGFPPDIDV 207

QY 161 EWESNGQ--PENNYKTTTPVLDSDGFFLYSKLTVDKSRQOQNVFSCVMHEALHNHYT 218  
Db 208 EWQRNGQPEPEGNRYRTTPQDDVDGTYFLYSKLAVDKASWQGGGIFQCAVMHEALHNHYT 267

QY 219 QKSLSPGK 228  
Db 268 QKSIFKTPGK 277

C:Superfamily: immunoglobulin C region; immunoglobulin homology  
F:133-202/Domain: immunoglobulin homology <IMM>

Query Match 64.1%; Score 917; DB 2; Length 328;  
Best Local Similarity 65.4%; Pred. No. 1.7e-22;  
Matches 166; Conservative 32; Mismatches 24; Indels 32; Gaps 4;

QY 3 KTHTC-----PPCP-----APELLGSPVFLFPKPKDTLMISR 36  
Db 79 KSYTCNVNHPATTTKVDKRVGTRKPPCPICPACESP-----GPSVFIFPPKPKDTLMISR 134

QY 37 TPEVTCVVVDVSHEDPEVKFNMYVDGVEVHNAKTKPREEQNSTYRVVSVLTVLHQDWL 96  
Db 135 TPQVTCVVVDVSENVDFVFNVDGVEVHTAQTREKEQFNSTYRVVSVLTVLHQDWL 194

QY 97 GREYCKVSNKALPAPIETISKAKGQPREPOVYTLPPSRDELTKNOVSLTCLVKGFP 156  
Db 195 GKEFKCKVNNKDLPAITRIISKAKGQTPREPOVYTLPPHAELSRKSVITCLVIGF 254

QY 157 DIAVESNGQ--PENNYKTTTPVLDSDGFFLYSKLTVDKSRQOQNVFSCVMHEALH 214  
Db 255 DIDVEQRNGQPEPEGNRYRTTPQDDVDGTYFLYSKFSVDKASWQGGGIFQCAVMHEALH 314

QY 215 NYHTOKSLSPGK 228  
Db 315 NYHTOKSLSPGK 328

RESULT 13  
I47162  
Ig gamma 4 chain constant region - pig (fragment)  
C:Species: Sus scrofa domestica (domestic pig)  
C:Date: 21-Feb-1997 #sequence\_revision 21-Feb-1997 #text\_change 21-Jan-2000  
C:Accession: I47162  
R:Kacskovics, I.; Sun, J.; Butler, J.E.  
J. Immunol. 153, 3565-3573, 1994  
A:Title: Five putative subclasses of swine IgG identified from the cDNA sequences of  
A:Reference number: I47158; MUID:95015845; PMID:7930579  
A:Accession: I47162  
A:Status: preliminary; translated from GB/EMBL/DBDJ  
A:Molecule type: mRNA  
A:Residues: 1-277 <KAC>  
A:Cross-references: EMBL:U03782; NID:g433129; PIDN:AAAS2220.1; PID:g433130  
C:Genetics:  
A:Gene: IgG4  
C:Superfamily: immunoglobulin C region; immunoglobulin homology  
F:82-151/Domain: immunoglobulin homology <IMM>

Query Match 63.7%; Score 911; DB 2; Length 277;  
Best Local Similarity 66.4%; Pred. No. 1.5e-22;  
Matches 166; Conservative 31; Mismatches 29; Indels 24; Gaps 4;

QY 3 KTHTC-----PPCP-----APELLG-SPVFLFPKPKDTLMISRTPEV 40  
Db 28 KSYTCNVNHPATTTKVDKRVGTRKPPCPICPACESP-----GPSVFIFPPKPKDTLMISRTPKV 87

QY 41 TCVVVDVSHEDPEVKFNMYVDGVEVHNAKTKPREEQNSTYRVVSVLTVLHQDWLNGKEY 100  
Db 88 TCVVVDVSENVDFVFNVDGVEVHTAQTREKEQFNSTYRVVSVLTVLHQDWLNGKEY 147

QY 101 KCKVSNKALPAPIETISKAKGQPREPOVYTLPPSRDELTKNOVSLTCLVKGFPDIAV 160  
Db 148 KCKVNNKDLPAITRIISKAKGQTPREPOVYTLPPTEELSRKSVITCLVIGFPPDIDV 207

QY 161 EWESNGQ--PENNYKTTTPVLDSDGFFLYSKLTVDKSRQOQNVFSCVMHEALHNHYT 218  
Db 208 EWQRNGQPEPEGNRYRTTPQDDVDGTYFLYSKLAVDKASWQGGGIFQCAVMHEALHNHYT 267

QY 219 QKSLSPGK 228  
Db 268 QKSIFKTPGK 277





GenCore version 5.1.4.p5.4578  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: April 21, 2003, 10:42:21 ; Search time 10.0941 Seconds  
(without alignments)  
1084.766 Million cell updates/sec

Title: 2LINK7LINK7

Perfect score: 1430

Sequence: 1 MDKTHTCPPCPAPELLGPPS.....AFDWLKAFYDKVAEKLKEAF 264

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.0

Searched: 112892 seqs, 41476328 residues

Total number of hits satisfying chosen parameters: 112892

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt\_40.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match %	Length	ID	Description
1	1233	86.2	330	1	GCL_HUMAN
2	1146	80.1	326	1	GC2_HUMAN
3	1142	79.9	327	1	GC4_HUMAN
4	1138	79.6	290	1	GC3_HUMAN
5	921	64.4	323	1	GC_RABIT
6	894	62.5	329	1	GC2_CAVPO
7	848	59.3	329	1	GC3_MOUSE
8	841	58.8	398	1	GC3M_MOUSE
9	840	58.7	333	1	GC3_RAT
10	827	57.8	393	1	GC1M_MOUSE
11	825	57.7	326	1	GC1_RAT
12	823	57.6	324	1	GC1_MOUSE
13	819	57.3	399	1	GC4M_MOUSE
14	818	57.2	329	1	GC2_RAT
15	813	56.9	330	1	GC3A_MOUSE
16	809	56.6	335	1	GC4A_MOUSE
17	793	55.5	322	1	GC4_RAT
18	791	55.3	405	1	GC5M_MOUSE
19	785	54.9	336	1	GC2_MOUSE
20	387	27.1	421	1	EPC_MOUSE
21	382	26.7	454	1	MUC_HUMAN
22	372	26.0	429	1	EPC_RAT
23	370	25.9	391	1	MUCB_HUMAN
24	370	25.9	455	1	MUC_MOUSE
25	369	25.8	428	1	EPC_HUMAN
26	369	25.8	476	1	MUCM_MOUSE
27	367	25.7	458	1	MUC_RABIT
28	366	25.6	479	1	MUCM_RABIT
29	356	24.9	454	1	MUC_MESAU
30	349	24.4	450	1	MUC_CANFA
31	347	24.3	457	1	MUC_SUNMU
32	336	23.5	4969	1	RYR2_RABIT
33	336	23.5	6486	1	TYCC_BACBR

#### RESULT 1

ID	GCL_HUMAN	STANDARD	PRT	330 AA
AC	P01857			
DT	21-JUL-1986 (Rel. 01, Created)			
DT	21-JUL-1986 (Rel. 01, Last sequence update)			
DT	16-OCT-2001 (Rel. 40, Last annotation update)			
DE	Ig gamma-1 chain C region.			
GN	IGHG1			
OS	Homo sapiens (Human)			
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
OC	Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.			
OX	NCBI_TaxID=9606;			
RP	SEQUENCE FROM N.A.			
RX	MEDLINE=82274238; PubMed=6287432;			
RA	Ellison J.W., Berson B.J., Hood L.E.;			
RT	"The nucleotide sequence of a human immunoglobulin C gamma1 gene."			
RL	Nucleic Acids Res. 10:4071-4079(1982).			
RN	[2]			
RP	SEQUENCE OF 1-135 (MYELOMA PROTEIN EU).			
RX	MEDLINE=71064024; PubMed=5489771;			
RA	Cunningham B.A., Rutishauser U., Gall W.E., Gottlieb P.D.,			
RT	Waxdal M.J., Edelman G.M.;			
RL	"The covalent structure of a human gamma G-immunoglobulin. VII. Amino acid sequence of heavy-chain cyanogen bromide fragments H1-H4."			
RN	Biochemistry 9:3161-3170(1970).			
RP	SEQUENCE OF 136-329 (EU).			
RX	MEDLINE=71064025; PubMed=5530842;			
RA	Rutishauser U., Cunningham B.A., Bennett C., Konigsberg W.H.,			
RT	Edelman G.M.;			
RL	"The covalent structure of a human gamma G-immunoglobulin. 8. Amino acid sequence of heavy-chain cyanogen bromide fragments H5-H7."			
RN	Biochemistry 9:3171-3181(1970).			
RP	SEQUENCE (MYELOMA PROTEIN NIE).			
RX	MEDLINE=77070269; PubMed=826475;			
RA	Ponstingl H., Hilschmann N.;			
RT	"The rule of antibody structure. The primary structure of a monoclonal IgG1 immunoglobulin (myeloma protein NIE). III. The chymotryptic peptides of the H-chain, alignment of the tryptic peptides and discussion of the complete structure."			
RL	Hoppe-Seyler's Z. Physiol. Chem. 357:1571-1604(1976).			
RN	[5]			
RP	SEQUENCE (MYELOMA PROTEIN KOL), AND DISULFIDE BONDS.			
RX	MEDLINE=83289131; PubMed=6884994;			
RA	Schmidt W.E., Jung H.-D., Palm W., Hilschmann N.;			
RT	"Three-dimensional structure determination of antibodies. Primary structure of crystallized monoclonal immunoglobulin IgG1 KOL, I."			
RL	Hoppe-Seyler's Z. Physiol. Chem. 364:713-747(1983).			
RN	[6]			
RP	DISULFIDE BONDS.			
RX	MEDLINE=71064027; PubMed=4923144;			
RA	Gall W.E., Edelman G.M.;			
RT	"The covalent structure of a human gamma G-immunoglobulin. X.			

Q92736 homo sapien  
P11716 oryctolagus  
P21817 homo sapien  
P01875 gallus gall  
P45443 neurospora  
P23098 tripneustes  
Q39565 chlamydomon  
P23085 heterodontu  
P20929 homo sapien  
P04114 homo sapien  
P39057 anthocidari  
Q39575 chlamydomon

#### ALIGNMENTS

RT Intrachain disulfide bonds.";  
 RL Biochemistry 9:3188-3196(1970).  
 RN [7]  
 RP DISULFIDE BONDS.  
 RX MEDLINE-77070267; PubMed-1002129;  
 RA Dreker L., Schwarz J., Reichel W., Hilschmann N.;  
 RA "Rule of antibody structure. The primary structure of a monoclonal  
 RT IgG1 immunoglobulin (myeloma protein Nie). I: Purification and  
 RT characterization of the protein, the L- and H-chains, the  
 RT cyanogen bromide cleavage products, and the disulfide bridges.";  
 RL Hoppe-Seyler's Z. Physiol. Chem. 357:1515-1540(1976).  
 RN [8]  
 RP X-RAY CRYSTALLOGRAPHY (2.9 ANGSTROMS).  
 RX MEDLINE-81208100; PubMed-7236608;  
 RA Deisenhofer J.;  
 RA "Crystallographic refinement and atomic models of a human Fc fragment  
 RT aureus at 2.9- and 2.8-A resolution.";  
 RL Biochemistry 20:2361-2370(1981).  
 CC -1- MISCELLANEOUS: NIE HAS THE GIM(17) ALLOTYPIC MARKER, 97-K, & THE  
 CC MARKER & THE GIM (NON-1) MARKERS.  
 CC -1- MISCELLANEOUS: NIE ALSO DIFFERS IN THE AMIDATION STATES OF  
 CC 35,116,198,269 & 272.  
 CC -1- MISCELLANEOUS: EU ALSO DIFFERS IN THE AMIDATION STATES OF RESIDUES  
 CC 155, 166, 177, 195, 198, 269, AND 272 AND IN THE ORDER OF RESIDUES  
 CC 268-272.  
 CC -1- MISCELLANEOUS: KOL ALSO DIFFERS IN THE AMIDATION STATES OF  
 CC RESIDUES 198,267&272.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 CC EMBL; J00228; AAC82527.1; ALT\_INIT.  
 DR PR; A02146; GHU.  
 DR PDB; 1FC1; 15-JUL-92.  
 DR PDB; 1FC2; 15-JUL-92.  
 DR Genew; HGNC:5525; IGHL1.  
 DR MIN; 147100; -.  
 DR InterPro; IPR003006; Ig\_MHC.  
 DR InterPro; IPR003597; Ig\_c1.  
 DR InterPro; IPR003600; Ig-like.  
 DR Pfam; PF00047; Ig; 3.  
 DR SMART; SM00410; Ig-like; 1.  
 DR SMART; SM00407; IGc1; 2.  
 DR PROSITE; PS00290; IG\_MHC; 2.  
 DR Immunoglobulin domain; Immunoglobulin C region; Glycoprotein;  
 KW 3D-structure.  
 FT NON\_TER 1 1  
 FT DOMAIN 1 98 CH1.  
 FT DOMAIN 99 110 HINGE.  
 FT DOMAIN 111 223 CH2.  
 FT DOMAIN 224 330 CH3.  
 FT DISULFID 27 83  
 FT DISULFID 103 103 INTERCHAIN (WITH LIGHT CHAIN).  
 FT DISULFID 109 109 INTERCHAIN (WITH HEAVY CHAIN).  
 FT DISULFID 112 112 INTERCHAIN (WITH HEAVY CHAIN).  
 FT DISULFID 144 204  
 FT DISULFID 250 308  
 FT CARBOHYD 180 180  
 FT MOD\_RES 330 330  
 FT VARIANT 97 97  
 FT VARIANT 239 239  
 FT VARIANT 241 241  
 FT VARIANT 241 241  
 FT STRAND 123 126

FT HELIX 130 134  
 FT TURN 136 137  
 FT STRAND 141 148  
 FT TURN 158 162  
 FT STRAND 163 164  
 FT TURN 165 166  
 FT STRAND 175 178  
 FT STRAND 183 190  
 FT HELIX 193 197  
 FT TURN 198 199  
 FT STRAND 202 206  
 FT STRAND 215 219  
 FT STRAND 227 227  
 FT STRAND 230 234  
 FT HELIX 238 240  
 FT TURN 241 242  
 FT STRAND 245 256  
 FT STRAND 260 266  
 FT TURN 267 268  
 FT STRAND 269 270  
 FT STRAND 274 276  
 FT TURN 280 281  
 FT TURN 283 284  
 FT STRAND 287 296  
 FT HELIX 297 301  
 FT TURN 302 303  
 FT STRAND 306 312  
 FT TURN 313 314  
 FT TURN 316 317  
 FT STRAND 320 324  
 SQ SEQUENCE 330 AA; 36106 MW; 3770EE106C2FA33D CRC64;  
 Query Match 86.2%; Score 1233; DB 1; Length 330;  
 Best Local Similarity 100.0%; Pred. No. 1.le-33;  
 Matches 227; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 2 DKTHTCPCPAPELLGGPSVFLFPPKPKDTLMISRTPEVTCVVYVDSHEDPEVKFNWYVD 61  
 DB 104 DKTHTCPCPAPELLGGPSVFLFPPKPKDTLMISRTPEVTCVVYVDSHEDPEVKFNWYVD 163  
 QY 62 GVEVHNAKTPREEQYNSTYRVYSLTLVLDHQLNGKEYCKVSNKALPAPIETISKAK 121  
 DB 164 GVEVHNAKTPREEQYNSTYRVYSLTLVLDHQLNGKEYCKVSNKALPAPIETISKAK 223  
 QY 122 GQPREPOVYTLPPSRDELTRKNQVSLTCLVKGFPYSDIAGEVESNGQPENNYKTTPPVLD 181  
 DB 224 GQPREPOVYTLPPSRDELTRKNQVSLTCLVKGFPYSDIAGEVESNGQPENNYKTTPPVLD 283  
 QY 182 DGSFFLYSKLTVDKSRWQQGNVSCFVSVMHEALHNHYTQKSLSLSPGK 228  
 DB 284 DGSFFLYSKLTVDKSRWQQGNVSCFVSVMHEALHNHYTQKSLSLSPGK 330.  
 RESULT 2  
 GC2\_HUMAN  
 ID GC2\_HUMAN STANDARD; PRT; 326 AA.  
 AC P01859;  
 DT 21-JUL-1986 (Rel. 01, Created)  
 DT 21-JUL-1986 (Rel. 01, Last sequence update)  
 DE 16-OCT-2001 (Rel. 40, Last annotation update)  
 GN IGHG2.  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 OX NCBI\_TaxID=9606;  
 RN [1]  
 RP SEQUENCE OF 2-326 FROM N.A.  
 RA MEDLINE-82197621; PubMed-6804948;  
 RA Ellison J.W., Hood L.E.;  
 RT "Linkage and sequence homology of two human immunoglobulin gamma  
 RT heavy chain constant region genes.";  
 RL Proc. Natl. Acad. Sci. U.S.A. 79:1984-1988(1982).



DR	Genew; HGNC:5526; IGHC2.
DR	MIM; 147110;
DR	InterPro; IPR003006; Ig_MHC.
DR	InterPro; IPR003597; Ig-cl.
DR	InterPro; IPR003600; Ig_like.
DR	Pfam; PF00047; ig; 3.
DR	SMART; SM00410; Ig_like; 1.
DR	SMART; SM00407; IGcl; 2.
DR	PROSITE; PS00290; IG_MHC; 2.
KW	Immunoglobulin domain; Immunoglobulin C region.
FT	NON_TER 1 1
FT	DOMAIN 1 98 CH1.
FT	DOMAIN 99 110 HINGE.
FT	DOMAIN 111 219 CH2.
FT	DOMAIN 220 326 CH3.
FT	DISULFID 14 14 INTERCHAIN (WITH A LIGHT CHAIN).
FT	DISULFID 27 83
FT	DISULFID 102 102 INTERCHAIN (WITH A HEAVY CHAIN).
FT	DISULFID 103 103 INTERCHAIN (WITH A HEAVY CHAIN).
FT	DISULFID 106 106 INTERCHAIN (WITH A HEAVY CHAIN).
FT	DISULFID 109 109 INTERCHAIN (WITH A HEAVY CHAIN).
FT	DISULFID 140 200
FT	DISULFID 246 304
FT	SITE 156 156
FT	MOD_RES 326 326
FT	VARIANT 60 60
FT	CONFLICT 109 109 AT OR NEAR THE COMPLEMENT-BINDING SITE.
FT	SEQUENCE 326 AA; 35884 MW; 8310678C687CF9C CRC64;
QY	Query Match 80.1%; Score 1146; DB 1; Length 326;
Db	Best Local Similarity 90.6%; Pred. No. 8.2e-31;
Db	Matches 212; Conservative 9; Mismatches 6; Indels 7; Gaps
QY	1 MDKT-----HTCCPCAPPELLGGPSVLEPPPKDKDTLMISRTPEVTCVVVDVSHEDPEV 54
Db	94 VDKTVERKCCVECPCCAPP--VAGSFVLFPKKPKDITMISRTPEVTCVVVDVSHEDPEV 154
QY	55 KFNHYVDGVEVHNNAKTPREQYNSTYRVVSVLTVLHQDWLNGLNGKEYCKVSNKALPAPIE 114
Db	153 QFNHYVDGVEVHNNAKTPREQOFNSTFRVSVLTIVQDWLNGLNGKEYCKVSNKGLPAPIE 211
QY	115 KTISKAKGPREPQVYTLPSPSRDELTKNQVSLTCLVKGFYPSDI AVESWGQPENNYKT 177
Db	213 KTISKTKGPREPQVYTLPSPSREEMTKNQVSLTCLVKGFYPSDI AVESWGQPENNYKT 273
QY	175 TppvldSGSFFLYSKLTVDKSRQQGVNFCSCVMHEALHNNHYTKQSLSLSPGK 228
Db	273 TPpWLDSGSFFLYSKLTVDKSRQQGVNFCSCVMHEALHNNHYTKQSLSLSPGK 326
RESULT 3	
GC4_HUMAN	STANDARD; PRT; 327 AA.
ID GC4_HUMAN	
AC P01861;	
DT 21-JUL-1986 (Rel. 01, Created)	
DT 21-JUL-1986 (Rel. 01, Last sequence update)	
DT 16-OCT-2001 (Rel. 40, Last annotation update)	
DE Ig gamma-4 chain C region.	
GN IGHG4.	
OS Homo sapiens (Human).	
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;	
OC Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.	
OX NCBI_TaxID=9606;	
RN [1]	
RP SEQUENCE FROM N.A.	
RX MEDLINE=83157104; PubMed=6299662;	
RA Ellison J.W., Buxbaum J.N., Hood L.E.;	
RT "Nucleotide sequence of a human immunoglobulin C gamma 4 gene.";	
RL DNA 1:11-18(1981).	
[2]	
RP SEQUENCE OF 1-30 AND 81-326.	
RX MEDLINE=70207560; PubMed=4192699;	







```
CC EMBL; J00451; NOT_ANNOTATED_CDS.
DR PIR; B02156; G3MSC.
DR HSSP; P01857; 1FC1.
DR InterPro; IPR003006; Ig_MHC.
DR InterPro; IPR003597; Ig_cl.
DR InterPro; IPR003600; Ig_like.
DR Pfam; PF00047; Ig; 3.
DR SMART; SM00410; Ig_like; 1.
DR SMART; SM00407; IGcl; 2.
DR PROSITE; PS00290; IG_MHC; 1.
KW Immunoglobulin domain; Immunoglobulin C region; Glycoprotein;
FT Transmembrane; Alternative splicing.
FT NON_TER 1 1
FT DOMAIN 98 113 CH1.
FT DOMAIN 114 223 HINGE.
FT DOMAIN 224 327 CH2.
FT TRANSMEM 345 362 CH3.
FT DOMAIN 363 398 POTENTIAL.
FT DOMAIN 333 342 CYTOPLASMIC (POTENTIAL).
FT CONFLICT 342 342 E -> G (IN REF. 2).
FT CONFLICT 388 388 E -> Q (IN REF. 2).
FT CONFLICT 388 388 P -> F (IN REF. 2).
SQ SEQUENCE 329 AA; 36228 MW; F45827174182BAD6 CRC64;

Query Match 59.3%; Score 848; DB 1; Length 329;
Best Local Similarity 60.1%; Pred. No. 7.3e-21;
Matches 152; Conservative 35; Mismatches 39; Indels 27; Gaps 2;

QY 3 KTHTC-----PP---CPAPELLGGPSVFLFPKPKDTLMIS 35
Db 77 QTVICNVHASPTELKRIEPRIPKSTPPGSSCPGNILGGPSVFIFFPKPKDALMIS 136
QY 36 RTEVTCVVVDVSHEDPEVKFNVDGVEVHNAKTKPRREQYNTYRVVSVLVVQHODWL 95
Db 137 LTPKVTCTVVVDVSHEDDPDVHVSFVFNKVEVHTAQTQPREAQYNSTFRVSALPIQHDWM 196
QY 96 NGREYKCKVSKALPAPIETKISKAGOPREPOVYITLPPSRDELTKNOVSLTCLVKGFP 155
Db 197 RGFEKCKVKNKALPAPIETKISKAGRAQTPQVYITPPPREQSKKVSCLTCLVNTFFS 256
QY 156 SDIAVESNGQPNKYKTPPVLDSGSEFLYSLKTVDKSRWQQGNVFCSCVMHEALHN 215
Db 257 EATSVVERNGEQDYKNTPIILSDSGTYFLYSLKTVDTDSWLGQEIFTCVSVVHEALHN 316
QY 216 HYTKQSLSLSPGK 228
Db 317 HHTQKNLSRSPGK 329

RESULT 8
GC3M_MOUSE STANDARD; PRT; 398 AA.
AC P03987;
DT 23-OCT-1986 (Rel. 02, Created)
DT 01-AUG-1991 (Rel. 19, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE Ig gamma-3 chain C region, membrane-bound form.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=85027161; PubMed=6092053;
RA Weis J.A., Word C.J., Rimm D., Der-Balan G.P., Martinez H.M.,
RA Tucker P.W., Blattner F.R.;
RT "Structural analysis of the murine IgG3 constant region gene.";
RL EMBO J. 3:2041-2046(1984).
RN [2]
RP SEQUENCE OF 328-398 FROM N.A.
RX MEDLINE=8401483; PubMed=6314258;
RA Komaromy M., Clayton L., Rogers J., Robertson S., Kettman J.,
RA Wall R.;
RT "The structure of the mouse immunoglobulin in gamma 3 membrane gene
segment.";
RL Nucleic Acids Res. 11:6775-6785(1983).
```

```
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; J00451; AAB59655.1;
DR EMBL; V01526; CAA24767.1; ALT_SEQ.
DR PIR; A02155; G3MSM.
DR HSSP; P01857; 1FC1.
DR InterPro; IPR003006; Ig_MHC.
DR InterPro; IPR003597; Ig_cl.
DR InterPro; IPR003600; Ig_like.
DR Pfam; PF00047; Ig; 3.
DR SMART; SM00410; Ig_like; 1.
DR SMART; SM00407; IGcl; 2.
DR PROSITE; PS00290; IG_MHC; 1.
KW Immunoglobulin domain; Immunoglobulin C region; Glycoprotein;
FT Transmembrane; Alternative splicing.
FT NON_TER 1 1
FT DOMAIN 98 113 CH1.
FT DOMAIN 114 223 HINGE.
FT DOMAIN 224 327 CH2.
FT TRANSMEM 345 362 CH3.
FT DOMAIN 363 398 POTENTIAL.
FT DOMAIN 333 342 CYTOPLASMIC (POTENTIAL).
FT CONFLICT 342 342 E -> G (IN REF. 2).
FT CONFLICT 388 388 E -> Q (IN REF. 2).
FT CONFLICT 388 388 P -> F (IN REF. 2).
SQ SEQUENCE 398 AA; 43929 MW; CF7E264B50A41B95 CRC64;

Query Match 58.8%; Score 841; DB 1; Length 398;
Best Local Similarity 55.0%; Pred. No. 2.3e-20;
Matches 153; Conservative 37; Mismatches 44; Indels 44; Gaps 3;

QY 8 PP---CPAPELLGGPSVFLFPKPKDTLMISRTPEVTCVVVDVSHEDPEVKFNKYVDGVE 64
Db 106 PGSSCPPGNILGGPSVFIFFPKPKDALMISLTPKVTCTVVVDVSEDDPDVHVSFVFNKE 165
QY 65 VHNAKTKPREEQYNTYRVVSVLVVQHODWLNKKEYCKVSKNKPALPAPIETKISKAGQP 124
Db 166 VHTAQTQPREAQYNSTFRVSALPIQHDWMRGKCKVNNKALPAPIETKISKPKGRA 225
QY 125 REPOVYITLPPSRDELTKNOVSLTCLVKGFPYPSDIAVESNGQPNKYKTPPVLDSGGS 184
Db 226 QTPQVYITPPPREQSKKVSCLTCLVNTFFSEATSVVERNGEQDYKNTPIILSDSGT 285
QY 185 FFYLSKLTVDKSRWQQGNVFCSCVMHEALHNHYTKQSLSPGKDWLKFYDKVAEKLKE 244
Db 286 YFLYSLKLTVDTSWLGQEIFTCVSVVHEALHNHHTQKNLSRSEPLE-----LNETCAE 337
QY 245 AFD-----WL 249
Db 338 AQDGLDGLWTITITIFISLFLSVCSASVTLFVKVWI 375

RESULT 9
GC3M_MOUSE STANDARD; PRT; 333 AA.
ID GCB_RAT
AC P20761;
DT 01-FEB-1991 (Rel. 17, Created)
DT 01-FEB-1991 (Rel. 17, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE Ig gamma-2B chain C region.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=89232738; PubMed=3149946;
```

```

RA Brueggemann M.;
RT "Evolution of the rat immunoglobulin gamma heavy-chain gene family.";
RL Gene 74:473-482(1988).
DR PIR: PS0018; PS0018.
DR HSP: P01842; 7FAB.
DR InterPro: IPR003006; Ig_MHC.
DR InterPro: IPR003597; Ig_c1.
DR InterPro: IPR003600; Ig_like.
DR Pfam: PF00047; Ig; 3.
DR SMART: SM00410; Ig_like; 1.
DR SMART: SM00407; Igc1; 2.
DR PROSITE: PS00290; Ig_MHC; 1.
KW Immunoglobulin domain; Immunoglobulin C region.
FT NON_TER 1
FT DISULFID 15 15 INTERCHAIN (WITH A LIGHT CHAIN).
FT DISULFID 27 80
FT DISULFID 106 106 INTERCHAIN (WITH A HEAVY CHAIN).
FT DISULFID 109 109 INTERCHAIN (WITH A HEAVY CHAIN).
FT DISULFID 112 112 INTERCHAIN (WITH A HEAVY CHAIN).
FT DISULFID 115 115 INTERCHAIN (WITH A HEAVY CHAIN).
FT DISULFID 147 207
FT DISULFID 253 311
SQ SEQUENCE 333 AA; 36497 MW; 55F8B64D48D460A6 CRC64;

Query Match 58.7%; Score 840; DB 1; Length 333;
Best Local Similarity 62.1%; Pred. No. 1.4e-20;
Matches 151; Conservative 34; Mismatches 43; Indels 15; Gaps 2;

QY 1 MKT-----HTCP-----PCPAPELLGGPSVFLPPPKDPTLMISRPETVCVV 45
DB 91 VDKVRRNGGIGKCPCTCHKCPVPELLGGPSVFLPPPKDPTLLISQNAKVCVV 150
QY 46 DYSHDEPVKNVYDGVVHNAKTPREEQYNSTYRVSVLTVLHQLDNLGCKEYCKV 105
DB 151 DVSEEDPVQFVFNNEVHTAQTPREEQNSTFRVVSALPIQHDMMSGKEPKCKVN 210
QY 106 NKALPAPIEKTISKAKGQPREPQVYTLPPSRDELATKNVSLTCLVKGFPSPDIATVWESN 165
DB 211 NKALPSPIEKTISKPKGLVKRPQVVMGPPTEQLTEQVSLTCLISGFLPNDIGVETSN 270
QY 166 GQPNENYKTPVLSDSGFFLYSKLTVDKSRWQGNVFCSCVMHEALHNHYTKSLSL 225
DB 271 GHIEKNYKTEPMDSDGFFMYSLNVERSDRAPFCVSVVHGLHNHVKESISRP 330
QY 226 PGK 228
DB 331 PGK 333

RESULT 10
GC1M_MOUSE
ID GC1M_MOUSE STANDARD; PRT; 393 AA.
AC P01869;
DT 21-JUL-1986 (Rel. 01, Created)
DT 01-AUG-1991 (Rel. 19, Last sequence update)
DT 30-MAY-2000 (Rel. 39, Last annotation update)
DE Ig gamma-1 chain C region, membrane-bound form.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=80045036; PubMed=115593;
RA Honjo T., Obata M., Yamawaki-Kataoka Y., Kataoka T., Kawakami T.,
RA Takahashi N., Mano Y.;
RT "Cloning and complete nucleotide sequence of mouse immunoglobulin
gamma 1 chain gene.";
RL Cell 18:559-568(1979).
[2]
RP SEQUENCE OF 323-393 FROM N.A.
RX MEDLINE=82197626; PubMed=6804950;
RA Tyler B.M., Cowman A.F., Gerondakis S.D., Adams J.M., Bernard O.;

```

```

RT "mRNA for surface immunoglobulin gamma chains encodes a highly
conserved transmembrane sequence and a 28-residue intracellular
domain.";
RL Proc. Natl. Acad. Sci. U.S.A. 79:2008-2012(1982).
RN [3]
RP SEQUENCE OF 323-366 FROM N.A.
RX MEDLINE=82115295; PubMed=679207;
RA Rogers J., Choi E., Souza L., Carter C., Word C.J., Kuehl M.,
RA Eisenberg D., Wall R.;
RT "Gene segments encoding transmembrane carboxyl termini of
immunoglobulin gamma chains.";
RL Cell 26:19-27(1981).
RN [4]
RP SEQUENCE OF 1-44 FROM N.A.
RX MEDLINE=8222190; PubMed=6283537;
RA Yamawaki-Kataoka Y., Nakai S., Miyata T., Honjo T.;
RT "Nucleotide sequences of gene segments encoding membrane domains of
immunoglobulin gamma chains.";
RL Proc. Natl. Acad. Sci. U.S.A. 79:2623-2627(1982).
CC -!- ALTERNATIVE PRODUCTS: CELL LINES PRODUCING IGG CONTAIN TWO MRNA
SPECIES FOR IG GAMMA CHAINS. THE MAJOR SPECIES ENCODES SECRETED
GAMMA CHAINS. A LESS ABUNDANT SPECIES APPEARS TO ENCODE MEMBRANE-
BOUND CHAINS IN THAT IT CONTAINS AN ALTERNATIVE 3' END, ENCODED
IN SEPARATE EXONS, THAT IS HOMOLOGOUS WITH THE MEMBRANE-BOUND
SEGMENT OF MU CHAINS.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch).
CC -----
DR EMBL; V00793; CAA24172.1; -
DR EMBL; V00793; CAA24173.1; -
DR EMBL; V00793; CAA24174.1; -
DR PIR; B02159; GLMSM.
DR HSP; P01842; 7FAB.
DR MGB; MGI:96446; Igh-4.
DR InterPro: IPR003006; Ig_MHC.
DR InterPro: IPR003597; Ig_c1.
DR Pfam; PF00047; Ig; 3.
DR SMART; SM00407; Igc1; 2.
DR PROSITE; PS00290; Ig_MHC; 1.
KW Immunoglobulin domain; Immunoglobulin C region; Glycoprotein;
KW Alternative splicing; Transmembrane.
FT NON_TER 1
FT DOMAIN 1 97 CH1.
FT DOMAIN 98 110 HINGE.
FT DOMAIN 111 217 CH2.
FT DOMAIN 218 324 CH3.
FT DISULFID 27 82
FT DISULFID 102 102 INTERCHAIN (WITH A LIGHT CHAIN).
FT DISULFID 104 104 INTERCHAIN (WITH A HEAVY CHAIN).
FT DISULFID 107 107 INTERCHAIN (WITH A HEAVY CHAIN).
FT DISULFID 109 109 INTERCHAIN (WITH A HEAVY CHAIN).
FT DISULFID 138 198
FT CARBOHYD 174 174 N-LINKED (GLCNAC. .).
FT DISULFID 244 302 POTENTIAL.
FT TRANSMEM 340 357 CYTOPLASMIC (POTENTIAL).
FT DOMAIN 358 393 CYTOPLASMIC 4387AICE27 CRC64;
SQ SEQUENCE 393 AA; 43386 MW; 4CC88343B7AICE27 CRC64;

Query Match 57.8%; Score 827; DB 1; Length 393;
Best Local Similarity 51.7%; Pred. No. 6.4e-20;
Matches 149; Conservative 49; Mismatches 46; Indels 44; Gaps 6;

QY 7 CPP--CPAPELLGGPSVFLPPPKDPTLMISRPETVCVVVDSHEDPEVFNWVDGVE 64
DB 104 CKPCICTVPEV---SSVFIFPPKPKDVLITLTLPKVTCTVVDISKDDPEVFSWFVDVE 160
QY 65 VHNARTKPREQYNSTYRVSVLTVLHQLDNLGCKEYCKVSNKALPAPIEKTISKAKGP 124

```

[illegible]

```
DR GlycoSuiteDB; P01868; -.
DR MGD; MGI:96446; Igh-4.
DR InterPro: IPR003006; Ig_MHC.
DR InterPro: IPR003597; Ig_cl.
DR Pfam: PF00047; Ig; 3.
DR SMART; SM00407; Igc1; 2.
DR PROSITE; PS00290; IG_MHC; 1.
KW Immunoglobulin domain; Immunoglobulin C region; Glycoprotein;
FT Alternative splicing.
FT NON_TER 1
FT DOMAIN 1 97
FT DOMAIN 98 110 CHL.
FT DOMAIN 111 217 HNGE.
FT DOMAIN 218 324 CH2.
FT DISULFID 27 82 CH3.
FT DISULFID 102 102 INTERCHAIN (WITH A LIGHT CHAIN).
FT DISULFID 104 104 INTERCHAIN (WITH A HEAVY CHAIN).
FT DISULFID 107 107 INTERCHAIN (WITH A HEAVY CHAIN).
FT DISULFID 109 109 INTERCHAIN (WITH A HEAVY CHAIN).
FT DISULFID 138 198 INTERCHAIN (WITH A HEAVY CHAIN).
FT CARBOHYD 174 174 N-LINKED (GLCNAC. .).
FT FTIG-CAR_000055.
FT DISULFID 244 302
FT MOD_RES 324 324 REMOVED POST-TRANSLATIONALLY.
FT CONFLICT 276 276 N -> D (IN REF. 3).
FT CONFLICT 278 278 N -> D (IN REF. 3).
SQ SEQUENCE 324 AA; 35704 MW; A338812F3D1F2C93 CRC64;

Query Match 57.6%; Score 823; DB 1; Length 324;
Best Local Similarity 51.3%; Pred. No. 4.8e-20;
Matches 142; Conservative 44; Mismatches 35; Indels 56; Gaps 3;

QY 5 HT-----CpP--CP 11
DB 51 HTPFAVLQSLDLYTLSSVTPSPRSEVTQNVHPASSTVKDKIVPRDCGCKPCIT 110
QY 12 APELLGGPSVFLFPPPKDLMISRTPEVTVVVDVSHEDPEVKFNWYVDGVEVHNATK 71
DB 111 VPEV---SSVFIPPKPKDVLITLTPKVTCTVVVDISKDPEVQFSWFVDVVEHTAQ 167
QY 72 PREEQYNSTYRVSVLTVLHODWLNGLKEYCKVSNKALPAPIETKISKAKGP 131
DB 168 PREEQNFSTFSVSELPIMHODWLNGLKEYCKVSNKALPAPIETKISKAKGP 227
QY 132 LPPSRDELTKNQVSLTCLVKGFYPSDIAVEWESNGQPENNYKTTPPVLDSDGSF 191
DB 228 IPPPKQMAKDKVSLTCMTDFEPDIIVIEWQWNGQPAENYKNTQPIIMNTNGSYFV 287
QY 192 TVDKSRWQGNVFCSCVMHEALHNHYTKLSLSLSPGK 228
DB 288 NVQKSNWEAGNFTCSVLHDEGLNHHTEKLSLSPGK 324

RESULT 13
GCAM_MOUSE
ID GCAM_MOUSE STANDARD; PRT; 399 AA.
AC P01865;
DT 21-JUL-1986 (Rel. 01, Created)
DT 01-AUG-1991 (Rel. 19, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE Ig gamma-2A chain C region, membrane-bound form.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=82222190; PubMed=6283537;
RA Yamawaki-Kataoka Y., Nakai S., Miyata T., Honjo T.;
RT "Nucleotide sequences of gene segments encoding membrane domains of
RL immunoglobulin gamma chains."
RL Proc. Natl. Acad. Sci. U.S.A. 79:2623-2627(1982).
CC -!- ALTERNATIVE PRODUCTS: CELL LINES PRODUCING IGG CONTAIN TWO MRNA
```

```
CC SPECIES FOR IG GAMMA CHAINS. THE MAJOR SPECIES ENCODES SECRETED
CC GAMMA CHAINS. A LESS ABUNDANT SPECIES APPEARS TO ENCODE MEMBRANE-
CC BOUND CHAINS IN THAT IT CONTAINS AN ALTERNATIVE 3' END, ENCODED
CC IN SEPARATE EXONS, THAT IS HOMOLOGOUS WITH THE MEMBRANE-BOUND
CC SEGMENT OF MU CHAINS.
CC -!- MISCELLANEOUS: THE SEQUENCE OF RESIDUES 1-329 IS ASSUMED TO BE
CC IDENTICAL WITH THE CORRESPONDING REGION OF THE SECRETED FORM OF
CC THE A ALLELE.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; J00471; AAB59661.1; ALT_INIT.
CC PIR; A02154; G2MSAM.
CC HSP; P01842; 7FAB.
CC MGD; MGI:96443; Igh-1.
CC InterPro: IPR003006; Ig_MHC.
CC InterPro: IPR003597; Ig_cl.
CC Pfam; PF00047; Ig; 2.
CC SMART; SM00410; Igc1; 2.
CC SMART; SM00407; Igc1; 2.
CC PROSITE; PS00290; IG_MHC; 1.
KW Immunoglobulin domain; Immunoglobulin C region; Glycoprotein;
KW Transmembrane; Alternative splicing.
FT NON_TER 1
FT DISULFID 15 15 INTERCHAIN (WITH A LIGHT CHAIN).
FT DISULFID 27 82
FT DISULFID 107 107 INTERCHAIN (WITH A HEAVY CHAIN).
FT DISULFID 110 110 INTERCHAIN (WITH A HEAVY CHAIN).
FT DISULFID 112 112 INTERCHAIN (WITH A HEAVY CHAIN).
FT DISULFID 144 204
FT DISULFID 250 308
FT TRANSMEM 346 363
FT DOMAIN 364 399
FT CARBOHYD 180 180
FT SEQUENCE 399 AA; 44020 MW; 4C38138BFAED3FF0 CRC64;

Query Match 57.3%; Score 819; DB 1; Length 399;
Best Local Similarity 54.3%; Pred. No. 1.2e-19;
Matches 152; Conservative 33; Mismatches 52; Indels 43; Gaps 4;

QY 7 CPP--CPAPELGGPSVFLFPPPKDLMISRTPEVTVVVDVSHEDPEVKFNWYVDGVE 64
DB 107 CPPCKCAPNLLGGPSVFLFPPPKDKVLMISLPIVTCVVVDVSDPDQVQISWFWNNVE 166
QY 65 VHNATKPREEQYNSTYRVSVLTVLHODWLNGLKEYCKVSNKALPAPIETKISKAKGP 124
DB 167 VHTAQTHREDYNSTLRVVSALPIQHDQWMSGKREKCKVNNKDLPAPIERTISKPKGSV 226
QY 125 REPQVYTLPPSRDELTKNQVSLTCLVKGFYPSDIAVEWESNGQPENNYKTTPPVLDSDGS 184
DB 227 RAPQVYVLPPEEEMTKKQVLTCTVDFMPEDIVVEWNTNGKTELNVKNTPEVLDSDGS 286
QY 185 FFLYSKLVTDKSRWQGNVFCSCVMHEALHNHYTKLSLSLSPGK----- 229
DB 287 YFMYSKLVREKKNVWERNYSVCSVVEHGLNHHHTKFSRTPGLDLDVCAEADQDGLD 346
QY 230 -W-----LKAFFDKVAEKLKE 244
DB 347 LWTITITIFISLISVSVCSASVTLFKVKWIFSSVVE-LKQ 385

RESULT 14
GCC_RAT
ID GCC_RAT STANDARD; PRT; 329 AA.
AC P20762;
DT 01-FEB-1991 (Rel. 17, Created)
```



```
DT 01-FEB-1991 (Rel. 17, Last sequence update)
DE 15-JUL-1999 (Rel. 38, Last annotation update)
DE Ig gamma-2C chain C region.
OC Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RN SEQUENCE FROM N.A.
RX MEDLINE=88166903; PubMed=3127222;
RA Bruggemann M., Delmastro-Galfre P., Waldmann H., Calabi F.;
RT "Sequence of a rat immunoglobulin gamma 2c heavy chain constant
RL Eur. J. Immunol. 18:317-319(1988).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: X07189; CAA30169.1; -.
DR PIR: S00847; S00847.
DR HSSP: P01842; 7FAB.
DR InterPro: IPR003006; Ig_MHC.
DR InterPro: IPR003597; Ig_cl.
DR InterPro: IPR003600; Ig_like.
DR Pfam: PF00047; Ig; 2.
DR SMART: SM00410; IG_Like; 1.
DR SMART: SM00407; IGcl; 2.
DR PROSITE: PS00290; IG_MHC; 1.
KW Immunoglobulin domain; Immunoglobulin C region.
FT NON_TER 1
FT DOMAIN 1 97 CH1.
FT DOMAIN 98 113 HINGE.
FT DOMAIN 114 222 CH2.
FT DOMAIN 223 329 CH3.
FT DISULFID 15 15 INTERCHAIN (WITH A LIGHT CHAIN).
FT DISULFID 27 82 INTERCHAIN (WITH A HEAVY CHAIN).
FT DISULFID 111 111 INTERCHAIN (WITH A HEAVY CHAIN).
FT DISULFID 113 113 INTERCHAIN (WITH A HEAVY CHAIN).
FT DISULFID 143 203 INTERCHAIN (WITH A HEAVY CHAIN).
FT DISULFID 249 307
SQ SEQUENCE 329 AA; 36571 MW; 5FCD7B7933850773 CRC64;

Query Match 57.2%; Score 818; DB 1; Length 329;
Best Local Similarity 57.9%; Pred. No. 7.3e-20;
Matches 146; Conservative 42; Mismatches 38; Indels 26; Gaps 2;

QY 3 KTHTC-----PP---CPAPELLGGPSVFLFPPKPKDTLMISR 36
Db 78 QTVCSVAHPATKSLNLIKRIEPRRPRTPTDTCSDNGLGVSFIFFPKPKDILMITL 137
QY 37 TPVETCVVDVSHEDPEVAFNVDGVGHNAKTKPREQYNSTYRVSVTLTVLHODWLN 96
Db 138 TPVETCVVDVSEEDVQSFNVDNVVFTACTQPHSQLNGTFRVSTLHLOHODWMS 197
QY 97 GKEYCKVSNKALPAPIETKISKAKQPREPVYTLPPSRDELTKNQVSLTCLVKGFYPS 156
Db 198 GKEFKCKVKNKOLPSPTEKTISKPRKARTPVYTIPTPPREQMKNKVSITCMVTSFYPA 257
QY 157 DIAVENESGCPENNYKTPPVLDSDGSEFLYSKLVNVDKSRWQGNVFCSSVHEALHNN 216
Db 258 SISVERNGELQDYKNTLPVLDSDSESYFLYSKLSVDSWNRGRDIYTCSSVHEALHNN 317
QY 217 YTKSLSLSPGK 228
Db 318 HTOKNLSRSPGK 329
RESULT 15
```

```
GCAA_MOUSE
ID GCAA_MOUSE STANDARD; PRT; 330 AA.
AC P01863;
DT 21-JUL-1986 (Rel. 01, Created)
DT 21-JUL-1986 (Rel. 01, Last sequence update)
DT 30-MAY-2000 (Rel. 39, Last annotation update)
DE Ig gamma-2A chain C region, A allele.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RN SEQUENCE FROM N.A.
RX MEDLINE=81076554; PubMed=6777755;
RA Sikorav J.-L., Auffray C., Rougeon F.;
RT "Structure of the constant and 3' untranslated regions of the murine
RL Balb/c gamma 2a heavy chain messenger RNA.";
RN [2]
RN Nucleic Acids Res. 8:3143-3155(1980).
RX SEQUENCE FROM N.A.
RX MEDLINE=81198976; PubMed=6262729;
RA Yamawaki-Kataoka Y., Miyata T., Honjo T.;
RT "The complete nucleotide sequence of mouse immunoglobulin gamma 2a gene
and evolution of heavy chain genes: further evidence for intervening
sequence-mediated domain transfer.";
RN [3]
RN Nucleic Acids Res. 9:1365-1381(1981).
RX SEQUENCE FROM N.A.
RX MEDLINE=81223894; PubMed=6787604;
RA Ollo R., Auffray C., Mochamps C., Rougeon F.;
RT "Comparison of mouse immunoglobulin gamma 2a and gamma 2b chain genes
RT suggests that exons can be exchanged between genes in a multigenic
family.";
RN [4]
RN Proc. Natl. Acad. Sci. U.S.A. 78:2442-2446(1981).
RX MYELOMA PROTEIN MOPC 173.
RX MEDLINE=74175517; PubMed=4831970;
RA Bourgois A., Fougereau M., Rocca-Serra J.;
RT "Determination of the primary structure of a mouse IgG2a
immunoglobulin:amino-acid sequence of the Fc fragment. Implications
for the evolution of immunoglobulin structure and function.";
RN [5]
RN Eur. J. Biochem. 43:423-435(1974).
RX DISULFIDE BONDS.
RX MEDLINE=73056887; PubMed=4565406;
RA de Preval C., Fougereau M.;
RT "Determination of the primary structure of a mouse gamma G2a
immunoglobulin. Identification of the disulfide bridges.";
RN Eur. J. Biochem. 30:452-462(1972).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: V00798; CAA24178.1; -.
DR PIR: A02152; G2MSA.
DR HSSP: P01842; 7FAB.
DR InterPro: IPR003006; Ig_MHC.
DR InterPro: IPR003597; Ig_cl.
DR InterPro: IPR003600; Ig_like.
DR Pfam: PF00047; Ig; 2.
DR SMART: SM00410; IG_Like; 1.
DR SMART: SM00407; IGcl; 2.
DR PROSITE: PS00290; IG_MHC; 1.
KW Immunoglobulin domain; Immunoglobulin C region.
FT NON_TER 1
FT DISULFID 15 15 INTERCHAIN (WITH A LIGHT CHAIN).
FT DISULFID 27 82 INTERCHAIN (WITH A HEAVY CHAIN).
FT DISULFID 107 107
```

FT	DISULFID	110	110	INTERCHAIN (WITH A HEAVY CHAIN).
FT	DISULFID	112	112	INTERCHAIN (WITH A HEAVY CHAIN).
FT	DISULFID	114	114	
FT	DISULFID	250	308	
FT	MOD.RES	330	330	REMOVED POST-TRANSLATIONALLY.
SQL	SEQUENCE	330 AA;	36389 MW;	B84361C5445A6864 CRC64;
	Query Match		56.9%;	Score 813; DB 1; Length 330;
	Best Local Similarity		52.9%;	Pred. No. 1.1e-19;
	Matches 148; Conservative		30; Mismatches 46;	Indels 56; Gaps 2;
QY	5 HT-----			-----CpP- 9
Db	51 HTFPAVLQSDLYTLSSSVTVTSSTPWSQSTCNVAHPASSSTKVKDKTKEPRGPTIKPCPC	110		
QY	10 -CPAPELLGGPSVLEPPPKDPTLMISRPEVTCVVVDVSHEDPEVKFNKYVDGVVHNA	68		
Db	111 KCPANLLGGPSVIFPPKIKDVLMLISLSPITVTCVVVDVSEDDPQVQISFVNNVVEHTA	170		
QY	69 KTKPREEQNSTVRVYSVLTVLHQDWLNGKEYKCKVSNKALPAPIEKTKSKAKGPREPQ	128		
	:			
Db	171 QQTQTHREDYNSTLRVVSALPIQHDHWMSGKEFKCKVNNKDLPAPIERTISKPKGSVRAPQ	230		
QY	129 VYTLPSRDELTKQNVSLTCLVKGFPYSDIAVESWNGQPPENNYKTTTPVLDSDGSEFLY	188		
Db	231 VYVLPPEEEMTKQVTLTCWVTDPEWDIYVEMTNGKTELNYKNTEPVLDSGGSFYMY	290		
QY	189 SKLTVDKSRWQQGNVFCSSVMHEALNNHYTKQSLSLSPGK	228		
Db	291 SKLVRKKKNWERNYSVCSVYVHEGLNHHHTKFSRTPGK	330		

Search completed: April 21, 2003, 10:43:13  
Job time : 11.0941 secs

GenCore version 5.1.4.p5.4578  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: April 21, 2003, 10:42:21 ; Search time 38.8235 Seconds  
(without alignments)  
1401.120 Million cell updates/sec

Title: 2LINK7LINK7

Perfect score: 1430

Sequence: 1 MDKTHTCPPCPAPPELLGGPS.....AFDWLKAFAVDKVAEKLKEAF 264

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.0

Searched: 671580 seqs, 206047115 residues

Total number of hits satisfying chosen parameters: 671580

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

SPTREMBL\_21:\*

1: sp\_archaea:\*

2: sp\_bacteria:\*

3: sp\_fungi:\*

4: sp\_human:\*

5: sp\_invertebrate:\*

6: sp\_mammal:\*

7: sp\_mhc:\*

8: sp\_organelle:\*

9: sp\_phase:\*

10: sp\_plant:\*

11: sp\_rodent:\*

12: sp\_virus:\*

13: sp\_vertebrate:\*

14: sp\_unclassified:\*

15: sp\_rvirus:\*

16: sp\_bacteriap:\*

17: sp\_archaeap:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1233	86.2	471	4	Q8TC77
2	1233	86.2	701	4	Q96FQ8
3	1143	79.9	473	4	Q8TC63
4	914	63.9	337	6	Q95M34
5	837	58.5	469	11	Q8R3V9
6	836	58.5	463	11	Q91LC4
7	831	58.1	437	11	Q9RIA4
8	816	57.1	473	11	Q9D8L4
9	811	56.7	473	11	Q99L25
10	810	56.6	468	11	Q99L31
11	788	55.1	474	11	Q8R3H6
12	787	55.0	473	11	Q91Z05
13	513	35.9	34350	4	Q8WZ42
14	493	34.5	26926	4	Q10466
15	493	34.5	26926	4	Q8WZB3
16	440	30.8	17352	5	Q95YM2

#### ALIGNMENTS

#### RESULT 1

Q8TC77 ID Q8TC77 PRELIMINARY; PRT; 471 AA.  
AC O8PC77;  
DT 01-JUN-2002 (TReMBLrel. 21, Created)  
DT 01-JUN-2002 (TReMBLrel. 21, Last sequence update)  
DE 01-JUN-2002 (TReMBLrel. 21, Last annotation update)  
DE Hypothetical 51.8 kDa protein.  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
OX NCBI\_TaxID=9606;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC TISSUE=SPLEEN;  
RA Strausberg R.;  
RL Submitted (FEB-2002) to the EMBL/GenBank/DBJ databases.  
DR EMBL; BC024289; AAH24289.1; -.  
KW Hypothetical protein.  
SQ SEQUENCE 471 AA; 3887F4CF588660E CRC64;

Query Match 86.2%; Score 1233; DB 4; Length 471;  
Best Local Similarity 100.0%; Pred. NO. 6.5e-39;  
Matches 227; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 2 DKTHTCPPCPAPPELLGGPSVFLFPPKPKDTLMISRTPEVTCVVDVSHEDPEVKFNWYVD 61  
Db 245 DKTHTCPPCPAPPELLGGPSVFLFPPKPKDTLMISRTPEVTCVVDVSHEDPEVKFNWYVD 304  
QY 62 GVEVHNAKTKPREQYNSTYRVSVLTIVLHODWLNKGEYKCKVSNKALPAPIETISKAK 121  
Db 305 GVEVHNAKTKPREQYNSTYRVSVLTIVLHODWLNKGEYKCKVSNKALPAPIETISKAK 364  
QY 122 GQPREPQYVTLPPSRDELTKNQVSLTCLVKGFYPSDIAVEWESNGQPENNYKTTTPVLDLS 181  
Db 365 GQPREPQYVTLPPSRDELTKNQVSLTCLVKGFYPSDIAVEWESNGQPENNYKTTTPVLDLS 424  
QY 182 DGSFFLYSKLTVDKSRWQQGNVSCVMHEALHNNHYTKQSLSPGK 228  
|||||

```
Db 425 DGSFFLYSKLTVDKSRWQGNVFCSCVMHEALHNHYTKQSLSLSPGK 471
RESULT 2
Q96PQ8 PRELIMINARY; PRT; 701 AA.
AC Q96PQ8;
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-MAR-2002 (TrEMBLrel. 20, Last annotation update)
DE Factor VII active site mutant immunocjugate.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=21477448; PubMed=11593034;
RA Hu Z., Garen A.;
RT "Targeting tissue factor on tumor vascular endothelial cells and tumor
RT cells for immunotherapy in mouse models of prostatic cancer.";
RL Proc. Natl. Acad. Sci. U.S.A. 98:12180-12185(2001).
DR EMBL: AF272774; AAK58686.1; -.
DR InterPro: IPR000152; ASX_hydroxyl.
DR InterPro: IPR000561; EGF-like.
DR InterPro: IPR000742; EGF_2.
DR InterPro: IPR001881; EGF_Ca.
DR InterPro: IPR003006; Ig_MHC.
DR InterPro: IPR001254; Ser.protease_Try.
DR InterPro: IPR000294; VitK_dep_GLA.
DR Pfam: PF00008; EGF; 2.
DR Pfam: PF00594; gla; 1.
DR Pfam: PF00047; ig; 2.
DR Pfam: PF00089; trypsin; 1.
DR SMART: SM00181; EGF; 2.
DR PROSITE: PS00010; ASX_HYDROXYL; UNKNOWN_1.
DR PROSITE: PS00022; EGF_1; UNKNOWN_1.
DR PROSITE: PS01186; EGF_2; UNKNOWN_1.
DR PROSITE: PS01187; EGF_Ca; UNKNOWN_1.
DR PROSITE: PS00011; GLU CARBOXYLATION; UNKNOWN_1.
DR PROSITE: PS00290; IG_MHC; UNKNOWN_1.
DR PROSITE: PS0240; TRYPSIN_DOM; 1.
DR PROSITE: PS00134; TRYPSIN_HIS; UNKNOWN_1.
DR PROSITE: PS00135; TRYPSIN_SER; UNKNOWN_1.
KW Hydrolase; Serine protease.
SQ SEQUENCE 701 AA; 77826 MW; 94AC6CEB42CC992F CRC64;

Query Match 86.2%; Score 1233; DB 4; Length 701;
Best Local Similarity 100.0%; Pred. No. 2.4e-38;
Matches 227; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 DKTHTCPCAPPELLGGPSVFLFPPKPKDTLMISRTPEVTCVVVDVSHEDPEVKFNWYVD 61
Db 475 DKTHTCPCAPPELLGGPSVFLFPPKPKDTLMISRTPEVTCVVVDVSHEDPEVKFNWYVD 534
QY 62 GVEVHNKATPREEQNSTYRVVSVLTVLHQDWLNGKEYKCKVSKNALKALPAIEKTSKAK 121
Db 535 GVEVHNKATPREEQNSTYRVVSVLTVLHQDWLNGKEYKCKVSKNALKALPAIEKTSKAK 594
QY 122 GQPREPQVYTLPPSRDELTKNQVSLTCLVKGFYPSDIAVEWESNGQPENNYKTPPPVLD 181
Db 595 GQPREPQVYTLPPSRDELTKNQVSLTCLVKGFYPSDIAVEWESNGQPENNYKTPPPVLD 654
QY 182 DGSFFLYSKLTVDKSRWQGNVFCSCVMHEALHNHYTKQSLSLSPGK 228
Db 655 DGSFFLYSKLTVDKSRWQGNVFCSCVMHEALHNHYTKQSLSLSPGK 701

RESULT 3
Q8TC63
ID Q8TC63 PRELIMINARY; PRT; 473 AA.
AC Q8TC63;
DT 01-JUN-2002 (TrEMBLrel. 21, Created)
```

```
DT 01-JUN-2002 (TrEMBLrel. 21, Last sequence update)
DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE Hypothetical 52.0 kDa protein.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=KIDNEY;
RA Strausberg R.;
RL Submitted (MAR-2002) to the EMBL/GenBank/DBJ databases.
RW EMBL; PC025985; AAK25985.1; -.
KW Hypothetical protein.
SQ SEQUENCE 473 AA; 51986 MW; E29920B09BA369F5 CRC64;

Query Match 79.9%; Score 1143; DB 4; Length 473;
Best Local Similarity 84.7%; Pred. No. 2.1e-35;
Matches 211; Conservative 9; Mismatches 6; Indels 23; Gaps 1;

QY 3 KTHT-----CPPCAPELLGGPSVFLFPPKPKDTLMISRTPE 39
Db 225 KYYTCNVNDHKPSNTKVDKRVESKYGPPCSCPAPEELGGPSVFLFPPKPKDTLMISRTPE 284
QY 40 VTCVVVDVSHEDPEVKFNWYVDGVEVHNKATPREEQNSTYRVVSVLTVLHQDWLNGKE 99
Db 285 VTCVVVDVSHEDPEVKFNWYVDGVEVHNKATPREEQNSTYRVVSVLTVLHQDWLNGKE 344
QY 100 YCKVSKNALKALPAIEKTSKAKGQPREPQVYTLPPSRDELTKNQVSLTCLVKGFYPSDIA 159
Db 345 YCKVSKNGLPSIEKTSKAKGQPREPQVYTLPPSRDELTKNQVSLTCLVKGFYPSDIA 404
QY 160 VEWESNGQPENNYKTPPPVLDSDGSFFLYSKLTVDKSRWQGNVFCSCVMHEALHNHYTQ 219
Db 405 VEWESNGQPENNYKTPPPVLDSDGSFFLYSKLTVDKSRWQGNVFCSCVMHEALHNHYTQ 464
QY 220 KSLSLSPGK 228
Db 465 KSLSLSLGK 473

RESULT 4
Q95M34 PRELIMINARY; PRT; 337 AA.
AC Q95M34;
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE Immunoglobulin gamma 1 heavy chain constant region
DE (Fragment).
GN IGHC1.
OS Equus caballus (Horse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Perissodactyla; Equidae; Equus.
OX NCBI_TaxID=9796;
RN [1]
RP SEQUENCE FROM N.A.
RA Wagner B.;
RL Submitted (DEC-2000) to the EMBL/GenBank/DBJ databases.
RN [2]
RX MEDLINE=98383416; PubMed=9717671;
RA Wagner B., Overesch G., Sheoran A., Holmes M., Richards C.,
RA Leibold W., Radbruch A.;
RT "Organization of the equine immunoglobulin heavy chain constant region
RT genes. III. Alignment of c-mu, c-gamma, c-epsilon and c-alpha genes.";
RL Immunobiology 199:105-119(1998).
DR EMBL: AJ300675; CAC44624.1; -.
DR InterPro: IPR003006; Ig_MHC.
DR Pfam: PF00047; ig; 2.
DR PROSITE: PS00290; IG_MHC; UNKNOWN_2.
FT NON_TER 1
SQ SEQUENCE 337 AA; 37438 MW; A60BF2B01DEFD1F6 CRC64;
```

```
Query Match 63.9%; Score 914; DB 6; Length 337;
Best Local Similarity 56.8%; Pred. No. 5.6e-27;
Matches 163; Conservative 32; Mismatches 29; Indels 63; Gaps 2;

QY 5 HT----- 6
DB 51 HTFPSVQLSSGFGSYLSMTVTPASTWTSETYICNVVHAASNFVKVRIPIPDNHQKQCD 110
QY 7 ---CPCPAPELLGGSVLFPKPKDITLMSRTPEVTCVVVDVSHEDPEVKFNWYVDGV 63
DB 111 MSKCPKCPAPELLGGSVLFPKPKDITLMSRTPEVTCVVVDVSHEDPEVKFNWYMDGV 170
QY 64 EVHNATKPREQYNSTYRVVSVLTVLHQDWLNGKEYKCKVSNKALPAPIEKTISKAKG 123
DB 171 EVTATRPKEQFNSTYRVVSVLTVLHQDWLNGKEYKCKVSNKALPAPIEKTISKAKG 230
QY 124 PREPQVYTLPPSDELTKNQVSTCLVKGFPYSDIAVEWESNGQP--ENNYKTTTPVLDS 181
DB 231 SQEPQVVLAPHDELKSKSVTVCLVKDFYPPEINIEWQSNQGOPELETYKYSTTQAQDS 290
QY 182 DGSFFLYSLKTVDKSRWQGNVFCSCVMHEALHNHYTQKSLSLSPGK 228
DB 291 DGSFFLYSLKTVDKSRWQGNVFCSCVMHEALHNHYTQKSLSLSPGK 337

RESULT 5
Q8R3V9 PRELIMINARY; PRT; 469 AA.
ID Q8R3V9;
AC Q8R3V9;
DT 01-JUN-2002 (TrEMBLrel. 21, Created)
DT 01-JUN-2002 (TrEMBLrel. 21, Last sequence update)
DE Hypothetical 52.0 kDa protein.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RA Strausberg R.;
RL Submitted (FEB-2001) to the EMBL/GenBank/DDBJ databases.
DR EMBL; BC024405; AAH24405.1; -
DR HSSP; P01842; 7FAB.
DR MGD; MGI:96446; Igh-4.
DR InterPro; IPR003597; Ig_c1.
DR InterPro; IPR003600; Ig_like.
DR InterPro; IPR003006; Ig_MHC.
DR InterPro; IPR003596; Ig_V.
DR Pfam; PF00047; Ig; 4.
DR SMART; SM00409; IG; 2.
DR SMART; SM00407; IGc1; 2.
DR SMART; SM00406; IGV; 1.
DR SMART; SM00410; IG_like; 1.
DR PROSITE; PS00290; IG_MHC; UNKNOW1.
SQ SEQUENCE 469 AA; 51976 MW; 534793F155D05457 CRC64;

Query Match 58.5%; Score 837; DB 11; Length 469;
Best Local Similarity 52.3%; Pred. No. 1.6e-23;
Matches 145; Conservative 41; Mismatches 35; Indels 56; Gaps 3;

QY 5 HT----- -CPP--CP 11
DB 196 HTFPAVQLSDLYTLSSSVTPSWPSTQVTCNVHPASSTKVDKIVPRDCGCKPCICT 255
QY 12 APPELLGSPVFLFPKPKDITLMSRTPEVTCVVVDVSHEDPEVKFNWYVDGVVHNAKTK 71
DB 256 VPEV---SSVFIFPPKPKDVLITLTPKVTCTVVDVSKDDPEVQFSFVDDVEVHTAQT 312
QY 72 PREQYNSTYRVVSVLTVLHQDWLNGKEYKCKVSNKALPAPIEKTISKAKGQPREPQVYT 131
DB 313 PREQFNSTFSVSELPIMHODWLNKGEKFCRVNSAFAFPAPIEKTISKGRKPAQVYT 372
QY 132 LPSPRDLTKNQVSLTCLVKGFYPSDIAVEWESNGQPENNYKTTTPPVLDSDGSFFLYSKL 191
DB 373 IPPPKQMAKDKVSLTCLMTITDFPEDITVEQWNGQPAENYKNTQIPMDTDSGYFYISKL 432
QY 192 TVDKSRWQGNVFCSCVMHEALHNHYTQKSLSLSPGK 228
DB 433 NVQSKNWEAGNTFTCSVLHGLNHHHTKSLSHSPGK 469

Query Match 58.5%; Score 836; DB 11; Length 463;
Best Local Similarity 52.0%; Pred. No. 1.7e-23;
Matches 144; Conservative 42; Mismatches 35; Indels 56; Gaps 3;

QY 5 HT----- -CPP--CP 11
DB 190 HTFPAVQLSDLYTLSSSVTPSWPSTQVTCNVHPASSTKVDKIVPRDCGCKPCICT 249
QY 12 APPELLGSPVFLFPKPKDITLMSRTPEVTCVVVDVSHEDPEVKFNWYVDGVVHNAKTK 71
DB 250 VPEV---SSVFIFPPKPKDVLITLTPKVTCTVVDVSKDDPEVQFSFVDDVEVHTAQT 306
QY 72 PREQYNSTYRVVSVLTVLHQDWLNGKEYKCKVSNKALPAPIEKTISKAKGQPREPQVYT 131
DB 307 PREQFNSTFSVSELPIMHODWLNKGEKFCRVNSAFAFPAPIEKTISKGRKPAQVYT 366
QY 132 LPSPRDLTKNQVSLTCLVKGFYPSDIAVEWESNGQPENNYKTTTPPVLDSDGSFFLYSKL 191
DB 367 IPPPKQMAKDKVSLTCLMTITDFPEDITVEQWNGQPAENYKNTQIPMDTDSGYFYISKL 426
QY 192 TVDKSRWQGNVFCSCVMHEALHNHYTQKSLSLSPGK 228
DB 427 NVQSKNWEAGNTFTCSVLHGLNHHHTKSLSHSPGK 463

RESULT 7
Q9RIA4 PRELIMINARY; PRT; 437 AA.
ID Q9RIA4;
AC Q9RIA4;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DE Gamma heavy chain of Mab7 (Fragment).
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RA Strausberg R.;
RL Submitted (FEB-2001) to the EMBL/GenBank/DDBJ databases.
DR EMBL; BC003439; AAH03435.1; -
DR HSSP; P01842; 7FAB.
DR MGD; MGI:96446; Igh-4.
DR InterPro; IPR003597; Ig_c1.
DR InterPro; IPR003600; Ig_like.
DR InterPro; IPR003006; Ig_MHC.
DR InterPro; IPR003596; Ig_V.
DR Pfam; PF00047; Ig; 4.
DR SMART; SM00409; IG; 2.
DR SMART; SM00407; IGc1; 2.
DR SMART; SM00406; IGV; 1.
DR SMART; SM00410; IG_like; 1.
DR PROSITE; PS00290; IG_MHC; UNKNOW1.
SQ SEQUENCE 463 AA; 51007 MW; EAA674C6BBC30783 CRC64;

Query Match 58.5%; Score 836; DB 11; Length 463;
Best Local Similarity 52.0%; Pred. No. 1.7e-23;
Matches 144; Conservative 42; Mismatches 35; Indels 56; Gaps 3;

QY 5 HT----- -CPP--CP 11
DB 190 HTFPAVQLSDLYTLSSSVTPSWPSTQVTCNVHPASSTKVDKIVPRDCGCKPCICT 249
QY 12 APPELLGSPVFLFPKPKDITLMSRTPEVTCVVVDVSHEDPEVKFNWYVDGVVHNAKTK 71
DB 250 VPEV---SSVFIFPPKPKDVLITLTPKVTCTVVDVSKDDPEVQFSFVDDVEVHTAQT 306
QY 72 PREQYNSTYRVVSVLTVLHQDWLNGKEYKCKVSNKALPAPIEKTISKAKGQPREPQVYT 131
DB 307 PREQFNSTFSVSELPIMHODWLNKGEKFCRVNSAFAFPAPIEKTISKGRKPAQVYT 366
QY 132 LPSPRDLTKNQVSLTCLVKGFYPSDIAVEWESNGQPENNYKTTTPPVLDSDGSFFLYSKL 191
DB 367 IPPPKQMAKDKVSLTCLMTITDFPEDITVEQWNGQPAENYKNTQIPMDTDSGYFYISKL 426
QY 192 TVDKSRWQGNVFCSCVMHEALHNHYTQKSLSLSPGK 228
DB 427 NVQSKNWEAGNTFTCSVLHGLNHHHTKSLSHSPGK 463
```

```

RT antibody (Wab 7, its light and heavy chains) and construction of a
RT single chain antibody (scFv).";
RL Submitted (MAY-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF152372; AAD40243.1; -.
DR HSSP; P01842; 7FAB.
DR MGD; MGI:96446; Igh-4.
DR InterPro; IPR003600; Ig_like.
DR InterPro; IPR003006; Ig_MHC.
DR InterPro; IPR003596; Ig_v.
DR Pfam; PF00047; Ig; 4.
DR SMART; SM00406; Igv; 1.
DR SMART; SM00410; Ig_like; 2.
DR PROSITE; PS00290; IG_MHC; UNKNOWN_1.
FT NON_TER 1
FT TER 437
SQ SEQUENCE 437 AA; 48142 MW; 5C3A7BB3EB7D697C CRC64;

Query Match 58.1%; Score 831; DB 11; Length 437;
Best Local Similarity 51.6%; Pred. No. 2.2e-23;
Matches 143; Conservative 43; Mismatches 35; Indels 56; Gaps 3;

Oy 5 HT-----CPP--CP 11
||
Db 164 HTFPAVLQSLDYLTLSSVTPSPSTWPTVCNVAHPASSTKVDKIVPRDCGCKPCICT 223
Oy 12 APELLGGPSVFLFPPKPKDTLMISRTPEVTCVVDVSHEDPEVKFNWYVDGVEVHNAKTK 71
||
Db 224 VPEV---SSVFIFFPKPKDLVLTITLPKVTCTVVDVSHEDPEVQSFVVDVVEVHTAQIQ 280
Oy 72 PREQYNSTRYRVSVLTVLHQDLWNGKEYKVKSNKALPAPIEKTISKAKGQPREPQVY 131
||
Db 281 PREQFNSTRVSELPIMHQDLWNGKEFKRVNSAAFPAPTEKTSKTRKPAQVYT 340
Oy 132 LPSPRDLTKNOVSLTCLVKGYFSPVDIAVEWESNGQPNENYKTTTPVLDSGDSFLYSKL 191
||
Db 341 IPPPEQMAKDKVSTCMITDFEPDITVEWMQNGPAENYKNTQPIMDTDGYSFYVSKL 400
Oy 192 TVDKSRWQOGNFGVSCVMHEALHNHYTKLSLSLSPGK 228
||
Db 401 NVQKSNWEAGNFTCSVLHLEGLHNHTKLSLSLSPGK 437

RESULT 8
Q9DBL4
ID Q9DBL4 PRELIMINARY; PRT; 473 AA.
AC Q9DBL4;
DT 01-JUN-2001 (TrEMBLrel. 17, Created)
DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE 1810060009rik protein.
GN IGH-1 OR 1810060009RIK.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=PANCREAS;
RX MEDLINE=21085660; PubMed=11217851;
RA Kawai J., Shinozawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,
RA Arakawa T., Hara A., Fukunishi Y., Konno H., Adachi J., Fukuda S.,
RA Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamanaka I.,
RA Saito T., Okazaki Y., Gojobori T., Bono H., Kasukawa T., Saito R.,
RA Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,
RA Fleischmann W., Gaasterland T., Gissi C., King B., Kochiwa H.,
RA Kuehl P., Lewis S., Matsuo Y., Nikaido I., Pesole G., Quackenbush J.,
RA Schriml L.M., Staubli F., Suzuki R., Tomita M., Wagner L., Washio T.,
RA Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barsh G.,
RA Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaldo M.F.,
RA Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.,
RA Gustincich S., Hill D., Hofmann M., Hume D.A., Kamiya M., Lee N.H.,
RA Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombaerts P.,
RA Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,

```

```

RA Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F.,
RA Suzuki H., Toyooka K., Wang K.H., Weitz C., Whittaker C., Wilming L.,
RA Wynshaw-Boris A., Yoshida K., Hasegawa Y., Kawaji H., Kohtsuki S.,
RA Hayashizaki Y.;
RT "Functional annotation of a full-length mouse cDNA collection.";
RL Nature 409:685-690(2001).
DR EMBL; AK007918; BAB25349.1; -.
DR HSSP; P01842; 7FAB.
DR MGD; MGI:96443; Igh-1.
DR InterPro; IPR003599; Ig.
DR InterPro; IPR003597; Ig-cl.
DR InterPro; IPR003600; Ig_like.
DR InterPro; IPR003006; Ig_MHC.
DR InterPro; IPR003596; Ig_v.
DR Pfam; PF00047; Ig; 4.
DR SMART; SM00409; IG; 2.
DR SMART; SM00407; IGcl; 3.
DR SMART; SM00406; IGV; 1.
DR SMART; SM00410; IG_like; 1.
DR PROSITE; PS00290; IG_MHC; UNKNOWN_1.
SQ SEQUENCE 473 AA; 51699 MW; 9DED57A514475FBB CRC64;

Query Match 57.1%; Score 816; DB 11; Length 473;
Best Local Similarity 50.5%; Pred. No. 1.1e-22;
Matches 144; Conservative 38; Mismatches 42; Indels 61; Gaps 1;

Oy 5 HT-----
||
Db 189 HTFPALLQSLYTLSSVTVTSNTWPSITTCNVAHPASSTKVDKIEPRVITQNPQPP 248
Oy 7 ---CPPCPAPELLGGPSVFLFPPKPKDTLMISRTPEVTCVVDVSHEDPEVKFNWYVDGV 63
||
Db 249 LKECPCAAPDLLGGPSVFIFFPKIKDVLMLSLSPWVTCVVDVSEDDPDVQISFVNVN 308
Oy 64 EVHNAKTRPEEQYNSTRYRVSVLTVLHQDLWNGKEYKVKSNKALPAPIEKTISKAKGQ 123
||
Db 309 EVHTAQCTHREDYNSTRYRVSVLTVLHQDLWNGKEYKVKSNKALPAPIEKTISKPRGP 368
Oy 124 PREQVYTLPPSRDELTKNOVSLTCLVKGYFSPVDIAVEWESNGQPNENYKTTTPVLDSG 183
||
Db 369 VRAPQVYVLPPEAEETKRFSLTCMITGFLPAELAVDWTNGRTEQYKNTATVLDSDG 428
Oy 184 SFFLYSKLTVDKSRWQOGNFGVSCVMHEALHNHYTKLSLSLSPGK 228
||
Db 429 SYFWYSKLRVQKSTWGRSGFLACSVVHGLHNHLTKTISRSLGK 473

RESULT 9
Q99L25
ID Q99L25 PRELIMINARY; PRT; 473 AA.
AC Q99L25;
DT 01-JUN-2001 (TrEMBLrel. 17, Created)
DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE Similar to RIKEN cDNA 1810060009 gene.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=;
RX Submitted (FEB-2001) to the EMBL/GenBank/DBJ databases.
RA EMBL; BC003888; AAH03888.1; -.
DR HSSP; P01842; 7FAB.
DR InterPro; IPR003599; Ig.
DR InterPro; IPR003597; Ig-cl.
DR InterPro; IPR003600; Ig_like.
DR InterPro; IPR003006; Ig_MHC.
DR InterPro; IPR003596; Ig_v.
DR Pfam; PF00047; Ig; 3.
DR SMART; SM00409; IG; 2.
DR SMART; SM00407; IGcl; 3.

```



```

RA Strausberg R.;
RL Submitted (JUL-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC010327; AAH10327.1;
DR MGD; MGI:2144967; AU044919.
DR InterPro; IPR000345; CytC_heme_bind.
DR InterPro; IPR003006; Ig_MHC.
DR Pfam; PF00047; Ig_3.
DR PROSITE; PS00190; CYTOCHROME_C; UNKNOWN_1.
DR PROSITE; PS00290; IG_MHC; UNKNOWN_1.
KW Hypothetical protein.
SQ SEQUENCE 473 AA; 51946 MW; CF625F008932AF12 CRC64;

Query Match 55.0%; Score 787; DB 11; Length 473;
Best Local Similarity 50.0%; Pred. No. 1.5e-21;
Matches 143; Conservative 34; Mismatches 47; Indels 62; Gaps 2;

QY 5 HT-----CP 8
II
Db 188 HTFPALLOGLYTMSSSVTPSPSTWPSQVTCTVAHPASSTVTDKKLEPSGPISTINPCP 247
QY 9 P-----CPAPELLGGPSVFLPDKPKDTLMISRTPEVTCVVVDVSHEDPEVKFNWYVDG 62
Db 248 PKECHCKCAPNLEGPSVFIIPNKKVLMISLTPKVTCTVVVDVSEDDPDVQISWVFN 307
QY 63 VEVHNAKTPREEQVNSTYRVSVLTVLHQDWLNGKEYKCKVSNKALPAPIEKTISKARG 122
Db 308 VEVHTAQTQTHREDYNSTIRVVSALPIQHDAMSGCKEPCVNNKDLPSPIERTISKIKG 367
QY 123 QREPOVYTLPSRDELTKNQVSLTCLVKGFYPSDIAVEWESNGQPENNYKTPPPVLSD 182
Db 368 LVRAPOVYTLPPAELQSLRDKVSLTCLVGFNPGDISVWTSNGHTENYKDTAPVLSD 427
QY 183 GSFELYSLKLTVDKSRWQGNVFCSCVMHEALHNHYTOKSLSLSPGK 228
Db 428 GSYFIYSLKLDITKWEKTDSDSCNVRHGLKNTYLLKTIKTSRSPGK 473

RESULT 13
Q8WZ42
ID Q8WZ42 PRELIMINARY; PRT: 34350 AA.
AC Q8WZ42;
DT 01-MAR-2002 (TremBLrel. 20, Created)
DT 01-MAR-2002 (TremBLrel. 20, Last sequence update)
DT 01-JUN-2002 (TremBLrel. 21, Last annotation update)
DE Titin.
GN TTN.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP MEDLINE=20309627; PubMed=10850961;
RA Freiburg A., Trombitas K., Hell W., Cazorla O., Fougereousse F.,
RA Cencner T., Kolmerer B., Witt C., Beckmann J.S., Gregorio C.C.,
RA Granzier H., Labeit S.;
RT "Series of exon-skipping events in the elastic spring region of titin
RT as the structural basis for myofibrillar elastic diversity.";
RL Circ. Res. 86:1114-1121(2000).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=21573839; PubMed=11717165;
RA Bang M.L., Centner T., Fornoff F., Geach A.J., Gotthardt M.,
RA McNabb M., Witt C.C., Labeit D., Gregorio C.C., Granzier H.,
RA Labeit S.;
RT "The complete gene sequence of titin, expression of an unusual ~700
RT kDa titin isoform and its interaction with obscurin identify a novel
RT Z-line to I-band linking system.";
RL Circ. Res. 89:1065-1072(2001).
DR EMBL; AJ277892; CAD12456.1;
DR InterPro; IPR000282; Cytok_receptor_2.
DR InterPro; IPR000719; Euk_pkinase.
DR InterPro; IPR000577; FG_Y_kin.

```

```

DR InterPro; IPR003961; FN_III.
DR InterPro; IPR001092; HLH_basic.
DR InterPro; IPR003599; Ig.
DR InterPro; IPR003598; Ig_c2.
DR InterPro; IPR003006; Ig_MHC.
DR InterPro; IPR003596; Ig_v.
DR InterPro; IPR002016; Peroxidase.
DR InterPro; IPR004168; PPAK_motif.
DR InterPro; IPR002290; Ser_thr_pkinase.
DR InterPro; IPR001245; Tyr_pkinase.
DR Pfam; PF00041; fn3; 132.
DR Pfam; PF00047; Ig; 146.
DR Pfam; PF00069; pkinase; 1.
DR Pfam; PF02818; PPAK; 53.
DR ProDom; PM000001; Euk_pkinase; 1.
DR SMART; SM00060; FN3; 133.
DR SMART; SM00409; IG; 167.
DR SMART; SM00408; IGC2; 148.
DR SMART; SM00406; IGV; 23.
DR SMART; SM00220; S_TKc; 1.
DR SMART; SM00219; Tyrc; 1.
DR PROSITE; PS00933; FG_Y_KINASES_1; UNKNOWN_1.
DR PROSITE; PS00038; HELIX_LOOP_HELIX; UNKNOWN_1.
DR PROSITE; PS00290; IG_MHC; UNKNOWN_1.
DR PROSITE; PS00435; PEROXIDASE_1; UNKNOWN_1.
DR PROSITE; PS00111; PROTEIN_KINASE_DOM; 1.
DR PROSITE; PS00109; PROTEIN_KINASE_TYR; UNKNOWN_1.
SQ SEQUENCE 34350 AA; 3816262 MW; 5B1120058A7CE58A CRC64;

Query Match 35.9%; Score 513; DB 4; Length 34350;
Best Local Similarity 0.5%; Pred. No. 7.4e-05;
Matches 172; Conservative 36; Mismatches 54; Indels 32039; Gaps 47;

QY 3 KTHCTCP-----9
Db 98 KAETAPPNFVRLQSMVTVRQGSQVRLQVRVTGIPTPVVKYRDGAIEOSSLDQFISQEGD 157
QY 10 -----9
Db 158 LYSLLIAEAYPEDSGTYSVNATNSVGRATSTAELLVQGEVEVPAAKTKTIYSTAQSISR 217
QY 10 -----9
Db 218 QTRIEKKIEAHFDARSATVEMVIDGAAGQQLPKHTPHRIPPKPSRSPTPSIAKAQL 277
QY 10 -----9
Db 278 ARQSPSPIRHSPSPVRHVRAPTPSPVRSVSPAARISTSPRSPLLMRKTQASTVAT 337
QY 10 -----9
Db 338 GPEVPPPKQEGYVASSSEAEEMRETTLTSTQIRTEWEGRYGVQEQVTISGAAGAAAS 397
QY 10 -----9
Db 398 VSASASAAEAATGATGKEVKQDADKAAAVATVVAADVMDARVREPVISAEQTAQTITTA 457
QY 10 -----9
Db 458 VHIQPAQEQVRKEAKTAVTKVVVAADKAKEQLKSRKTEVITTKQEQMHVTHEQIRKET 517
QY 10 -----9
Db 518 EKTFFVKVVISAAKAKEQETRISEETKKQKVQTEAIRQETITEAASVWVVAATAKTKL 577
QY 10 -----9
Db 578 ETVPGAQEBETTQDDQMHLSYEKIMKTRKTVVPRKIVATPKVRQDLVSRGREGITTKR 637
QY 10 -----9
Db 638 EQVITQEKMRKEAKETALTALSTIAVATAKAKEQETILTRPRETMATREQEIQVTHGKVDVGK 697

```



QY 10 ----- 9  
Db 698 KAEAVTVAADQAVREPREPHLEESVAQOTTLEYGYKERISAAKVAEPPQRPASEP 757  
QY 10 ----- 9  
Db 758 HVPKAVKPRVIOAPSETHIKTTQKGMHISQIKKTTDLTTERLVHVDKRPRTASPHFT 817  
QY 10 ----- 9  
Db 818 VSKISVPKTEHGEYASAGSAIATLOKELSATSSAQKITKSVKAPTVPKPSRTRVRAEPTP 877  
QY 10 ----- CPAP-- 13  
Db 878 LPOFPADTPTYKSEAGVEVKEVGSITGTVREBERFEVLHGREAKVTETARVPAPVE 937  
QY 14 ----- ELGGPSYFL-----FP----- 25  
Db 938 IPVTPPTLVGLKNVTIEGESVTLECHISGYPSPTVTVYREDYQIESSIDFQITFQSGI 997  
QY 26 ----- 25  
Db 998 ARLMIREAFADSGRTCSAVNEAGIVSTSCYLAVQVSEFEKETTAVTEKFTTEKRFV 1057  
QY 26 ----- PKP----- 28  
Db 1058 ESRDVVMTDTSLEQAGPGEAPAPVFTKPVVOKLVEGGSVVFQVGGNPKPHVYWK 1117  
QY 29 ----- 28  
Db 1118 SGVPLTGYRYKVSYNKOTGECKLVISMFTFADDAGEYTIIVRNKHGETSASALLEEADY 1177  
QY 29 ----- KDTLM-- 33  
Db 1178 ELLMKSOEMLYQVTAFOEPKVGSETAPGFVYSEYEKEQALIRKKMAKDITVVR 1237  
QY 34 ----- 33  
Db 1238 TYVEDQEFHISPEERLIKEIYRIIKTTLEELLEEDGEKMAVDISEASEVSGFDLRI 1297  
QY 34 ----- 33  
Db 1298 KNYRILEGMGVTHCKMSGYPLPKIAYKDGKRIKHGERYQMDFLODGRASLRIPVVLPE 1357  
QY 34 ----- ISRTP-- 38  
Db 1358 DEGIYAFASNIKNAICSGKLYVEPAAPLGAFTYIPTLEPVSRIKSLSPRSVSRPIRM 1417  
QY 39 ----- 38  
Db 1418 SPARMSPARMSPARMSPARMPCGRLEETDESQERLYKPVFVLKPVSKFCLEGQTARFD 1477  
QY 39 ----- 38  
Db 1478 LKVVGRMPETFWFHDGQIVNDYTHKVVIKEDGTQSLIIVPATPSDGEWTVVAQNRA 1537  
QY 39 ----- 38  
Db 1538 RSSISVILTVEAVHQVPMFVEKLNVNKEGSQLKMKVRATGNPNPDIVWLKNSDIIV 1597  
QY 39 ----- 38  
Db 1598 PHYPKIRIEGTGEAALKIDSTVSQDSAWYTATAINKAGRDITRCKVNVVEFAPEPE 1657  
QY 39 ----- 38  
Db 1658 RKLIIIPRGTYRAKEIAAEPLHLRYGOEWEEDLYDKEKQKPPFKKLTSLRKRF 1717  
QY 39 ----- EUTC 42  
Db 1718 GPAHFECRLTPIGDTMVVWHLHDGKPLEANRLRMINFEGYCSLDYGVAYSRDSGIITC 1777  
QY 43 ----- 42

Db 1778 RATNKYGTDHTSATLIVKDEKSLVEESQLPEGRKGLQRIEBELERMAHEGALTGVTTDOKE 1837  
QY 43 ----- 42  
Db 1838 KQKPDIVLYPEPVRVLEGETARFCRVGTGYPQKVNMYLNGQLIRKSRFRVRYDGIHYL 1897  
QY 43 ----- 42  
Db 1898 DIVDCKSYDTGEVKTVAENPEGVIEHKVKLEIQOQREDFRSVLRAPRPEPFVHPEGKL 1957  
QY 43 ----- 42  
Db 1958 QFEVQKVDROPVDTTETKEVVKLKRAERITHEKVPSESEELRSKFRTEGYEAITAVE 2017  
QY 43 ----- 42  
Db 2018 LKSRKXDESVEELLRTKDELLHWTKELTTEKKALAEKGKITPTPKPKIELSPSMEA 2077  
QY 43 ----- V 43  
Db 2078 PKIFERIOSQTVGGSDAHEFRVRVVGKPDPECEWYKNGVKIERSDRIYVWPNVCELV 2137  
QY 44 VVDVSHED----- 51  
Db 2138 IRDVTAEADSASIMVKAINIAGETSSHAFLVLQAKQLITFTQELQDVVAKEDTMAFECE 2197  
QY 52 ---PEVKFNWYDGVVH-----NAKT- 70  
Db 2198 TSEPFVKVWKYKDGMEVHEGDKYRMHSDRKHVFLSILTIDTDAEDYSCVLVEDENVKTT 2257  
QY 71 ----- 70  
Db 2258 AKLIVEGAVVEFVKELQDIEVPESYSELECIYSPENIEGKWHNDVELKSNKYTTISR 2317  
QY 71 ----- KPR----- 73  
Db 2318 RGQNLTVKDVTKEDQGEYSFVIDGKTKCKLKMKPRPIALQGLSDQKVCBGDIVQLEV 2377  
QY 74 ----- 73  
Db 2378 KVSLESVEGVMKDGQEVQPSDRVHIVIDKOSHMLLIEDMTKEDAGNYSFTIPALGLSTS 2437  
QY 74 ----- 73  
Db 2438 GRVSVYSDVITPLKDVNVIEGTAVLECKVSPDVTSVKWLNDEQIKPDDRQVAIVKG 2497  
QY 74 ----- 73  
Db 2498 TKQRLVINRTHASDEGPYKLIIVGRVETNCNLSVEKIKIIRGLRDLTCTETQNVVFEVELS 2557  
QY 74 ----- 73  
Db 2558 HSGIDVLWNFKDKIEKPSKSKYKIEAHGKIYKLVNMMKDDGKYTFYAGENITSGKLTV 2617  
QY 74 ----- 73  
Db 2618 AGGAISKPLTDQTVAESQEAFFCEVANPDSKGEWLBDGKHLPLTNINIRSESDGKRRLI 2677  
QY 74 ----- 73  
Db 2678 IAATKLDDIGEYTYKVATSKTSAKLKVEAVKIKKTLKNLTVTETQDAVFTVELTHPNVKG 2737  
QY 74 ----- 73  
Db 2738 VOWIKNGVLESNEKYAISVKGTIYSLRIKNCIAIVDESIFYFRGLRGLGASARLHVETKI 2797  
QY 74 ----- 73  
Db 2798 IKPKDVTALENATVAFESVSHDTPVVKWFHKNVEIKPSDKHRLVSRKVKHMLQNIS 2857  
QY 74 ----- 73

Db 2858 PSDAGEYTA VVGQLECKAKLFVETLHITKMTKNIEVPETKTASFCEVSHFNVPMSWLKN 2917  
Qy 74 ----- 73  
Db 2918 GVEIEMSEKFI VVGQKLHQLIIMNTSEDSAEYTFVCGNDQVSATLVTPIIMITSMKLD 2977  
Qy 74 ----- 73  
Db 2978 INAEKDIRITFEVTVYEGISYKWLKNGVEIKSTDKCOMRTKKLTHSLNIRNVHFGDAAD 3037  
Qy 74 ----- 73  
Db 3038 YTFVAGKATSTATLYVEARHIEFRKHIDIKVLEKKRAMFECEVSEPDITVQMMKDDQEL 3097  
Qy 74 ----- 73  
Db 3098 QITDRIKIQEKYVHURLLIPSTRMSDAGKYTVVAGGNVSTAKLFVEGRDVRIRSIKKEVO 3157  
Qy 74 ----- 73  
Db 3158 VIEKQRAVVEFEVNEDDVAHWYKDGIEINFQVQERHKYVVERRIHRMFISETRQSDAGE 3217  
Qy 74 ----- -BEQ 76  
Db 3218 YTFVAGNRSSVTLVYNAPEPQVLQELQPVTVQSGKPARFCAVISGRPQPKISWYKEEQ 3277  
Qy 77 YNST----- 80  
Db 3278 LLSTGFKCKFLHDGOEYTLILLIEAFPDAAYTCEAKNDYGVATTISASLSVEPVSFDP 3337  
Qy 81 ----- 80  
Db 3338 QEMPVYPAIITPLQDVTVTSEGPAPFQCRVSGTDLKVSWYSKDKKIKPSRFRMTQPED 3397  
Qy 81 ----- 80  
Db 3398 TYQLEIAEYPEDEGTYTFVNASNAVQVSVSTANLSLEAPESILHERIQEIEMEMKEPFSS 3457  
Qy 81 ----- 80  
Db 3458 SFLSAEEBGLHSAELQLSKINETLELLESPIVSTKFDSEKEGTCPIFIKEVSNADISMG 3517  
Qy 81 ----- 80  
Db 3518 DVATLSVTVIGIPKPIQWFFNGVLLTPSADYKFVFGDDHSLIILFTKLEDEGEYTCMA 3577  
Qy 81 ----- 80  
Db 3578 SNDYKGTICSAYLKINSKGEGHKDTETESAVAKSLEKLGGPCPPHFLKELKPIRCAQGLP 3637  
Qy 81 ----- 80  
Db 3638 AIFEYTVVGEPAPTVTWFENKQLCTSVYIIHNPNSSGTFIVNDPQREDSGLYICKAE 3697  
Qy 81 ----- 80  
Db 3698 NMLGESTCAELLVLEDMDTDPCKAKSTPEAPDPPTQPLKPAVEALDSEQIATF 3757  
Qy 81 ----- 80  
Db 3758 VKDTILKAALITEENQQLSYEHIAKANELSSQLPLGAQELQSLQKLTPESTREFLCI 3817  
Qy 81 ----- 80  
Db 3818 NGSIHFQPLKEPSPNLQIIVQSQKTFSEKIGILMPEEPETQAVLSDTEKIFPSAMSIEQI 3877  
Qy 81 ----- 80  
Db 3878 NSLTVPLKTLAABEPGPNQPOSSIEPPMHSYLTSVAEEVSPKKEKTVSDTNREQRVTLOK 3937  
Qy 81 ----- 80  
Db 3938 QEAQSALISLSLAEGHVESLQSPDVMISQVNYEPLVPSEHSCFEGGKILIESANPLENA 3997

Qy 81 ----- 80  
Db 3998 QODSAVRIEKGSLRFPPLALEEKQVLLKEEHSNDNVMPDPQIIIESKREPVAIKKVOEQVQ 4057  
Qy 81 ----- 80  
Db 4058 RDLLESKESLLSGIPPEQRNLNKIQICRALQAASEQPLFSEWLNRIEKEVEAVNITQ 4117  
Qy 81 ----- 80  
Db 4118 EPRHIMCYLVTSAKSVTEVTIIIEDVDPOMANLKMELRDALCAIIEEIDILTAEGPR 4177  
Qy 81 ----- 80  
Db 4178 IQQAKTSLQEBMDSFSGSQKVEPITEPEVESKYLISPEEVSYNVQSRVKYLDATPVTK 4237  
Qy 81 ----- 80  
Db 4238 GVASAVSDEKQESLKPSEKESSBSGTEEVATVKIQEAGGFIKEDGPMIHTPLVD 4297  
Qy 81 ----- 80  
Db 4298 TVSEEGDIVHLTTSITNAKEVNWYFENKLVPSDEKFKCLQDONTYTLVIDKVNTEHDQGE 4357  
Qy 81 ----- 80  
Db 4358 YVCEALNDSGKTATSAKLTVVKRAAPVIKRIEPLVHALGHAKFTCEIQSNPNVRFOWF 4417  
Qy 81 ----- 80  
Db 4418 KAGREIVESDKCSIRSSKYISSLEILRTQVVVDCGEYTCCKASNEYGSVCTATLTVTAEYP 4477  
Qy 81 ----- 80  
Db 4478 PTFLSRPKSLTTFVGKAAKFTCTVTGTPVETIWKOGAALSPSPNWKISDAENKHILEL 4537  
Qy 81 ----- 80  
Db 4538 SNLTIQDRGVYSCASNKFGADICQAEI IIDKPHFIKEPQVQSAINKKVHLECOVDDE 4597  
Qy 81 ----- 80  
Db 4598 RKVTWTWSKDGKLPQPKDYKICFEDKIATLEIPLAKLKDSTYVCTASNEAGSSCSAT 4657  
Qy 81 ----- 80  
Db 4658 VTVREPPSFVKKVDPSYLMPLPGESARLHCKLKGSPVIQVTFWPKNNKELSESNTVRMYEYN 4717  
Qy 81 ----- 80  
Db 4718 SEAILDITDKVVEDSGSYSCAVNDVGSDSCSTEIVIKEPPSFIKTLEPADIVRGTNALL 4777  
Qy 81 ----- 80  
Db 4778 QCEYSGTGPFPEISWFKDKQIRSSKKYRLFQSKSLVCLEIFSPNSADVGEYCVVANEVG 4837  
Qy 81 ----- 80  
Db 4838 KCGCMATHLLKEPPTFFVKVDDLLALGGQVTVLQAAVRGSEPISVTMWKGQEVIREDGKI 4897  
Qy 81 ----- 80  
Db 4898 KMSFSGVAVILLIPDVOISFGGKYTKLAENEAGSQTSGVGLIVKEPAKIIERAELIOVTA 4957  
Qy 81 ----- 80  
Db 4958 GDPATLEYTVAGTPELPKPKWYKDGRLVASKKYRISFPKNVNAQLKFFYSAELHDSQGTTFE 5017  
Qy 81 ----- 80  
Db 5018 ISNEVGSSCETTTTVLORDIAPFTTKPLRNVDSVWNGTCRLDCKIAGSLPMRVSNFKDG 5077

QY	81	-----80
Db	5078	KEAASDRYRIAFVEGTASLEIIRVDMNDAGNETCRATNSVGSKDSSGALIVQEPSPFT 5137
QY	81	-----80
Db	5138	KPGSKDVLPGSACLKSTFGOSTPLTIRWPKGNKELVSGSCYITKEALESSELYLVKT 5197
QY	81	-----80
Db	5198	SDSGTYTCKVSNVAGGYECSANLFFVEKLEPQSOLLKKGDATQLACKVTGTPPIK 5257
QY	81	-----80
Db	5258	ITWFANDREIKESKHRMSFVESTAVLRLTDVGIEDSGEYMCBAQNEAGSDHCSSIVIVK 5317
QY	81	-----80
Db	5318	ESPYFTKEFKPIEVLKEYDVMLLAEVAGTPPFEITWFKDNTILRSGRKYKTFIQDHLVSL 5377
QY	81	-----80
Db	5378	QILKFVAADAGEYQCRVTNEVGSSICARSRTLREPPSFIKKIBESTSSLRGTAFAQTALK 5437
QY	81	-----80
Db	5438	GSLPITVTWLKDSDEITDDNIRMTFENNVAASLYLSGIEVKHDGKYVCAKNDAGIORCS 5497
QY	81	-----80
Db	5498	ALLSVKEPATITEEAVSIDVTQGDPAFLQVFKSGTEKAITAKWFKDGOELTLGSKYKISVT 5557
QY	81	-----80
Db	5558	DTVSIKLIISTEKKDSGEYTFEVDNDVGRSCKARINVLIIIPPSFTKKLKMDSIKGS 5617
QY	81	-----80
Db	5618	FIDLECIVAGSHPIQWFKDDQEIASEKYKFSFHDTAFLEISOLEGTDGTYTCSAT 5677
QY	81	-----80
Db	5678	NKAGHNOCSGHLTVKEPPYFEVERQSDVNPNTRVOLKALVGGTAPMTIKWFKDNKELHS 5737
QY	81	-----80
Db	5738	GAARSVKDDTSTSLFPAKATDSGTYICQLSNDVGTATSKATLFWKEPPQFIKKPSPV 5797
QY	81	-----80
Db	5798	LVLNRGOSTTFEQITGTPKIRVSWYLDGNEITAIOKHGISFIDGLATFOISGARVENSG 5857
QY	81	-----80
Db	5858	TYVCEARNDAGTASCIELKVKEPPTFIRELKPVEVVKYSDVELECEVTGTPPFEVTLK 5917
QY	81	-----80
Db	5918	NNREIRSSKKYTLTDRVSVFNHLITKCDPSDTGEYQCIVSNEGSCSCTRVALKEPPSF 5977
QY	81	-----80
Db	5978	IKKIENTTVLKSSATFQSTVAGSPPISTWLKDDQILDEDDNVYISFVDSVATLQIRSV 6037
QY	81	-----80
Db	6038	DNGHSGRYTCAKNESGVERCYAFLVQEPQAIVERAKSVDTVTEKDPMTLECVVAGTPEL 6097
QY	81	-----80
Db	6098	KVKWLKDGKQIVPSRPSMSENNVASFRIQSVWVKQDSGOYTFKVENDFGSSCDAYLRV 6157
QY	81	-----YRVV---84

Db	6158	LDQNIPPSFTTKLTKMDKVLGSSIHMECKVSGSLPISAQWFKDGKEISTSAKYRLVUCHER 6217
QY	85	-----84
Db	6218	SVSLEVNNELEDTANTYCKVSNVAGDDACSGILTVKEPPSFLVKPGRQQAIPDSTVEFK 6277
QY	85	-----84
Db	6278	AILKGTPPFKIKWFKDDVELVSGPKCFIGLEGSTSFNLNLYSVDASKTCQYTCYHTNDVGS 6337
QY	85	-----84
Db	6338	DSCTTMLLVTEPPKFKVKLEASKIVKAGDSRLECKIAGSPEIRVWVFRNEHELPAUSDY 6397
QY	85	-----84
Db	6398	RMTFIDSVAVIQMNNLSTEDSGDFICEAQNPAGSTSCSTKIVVKEPPVFSFPPIVETLK 6457
QY	85	-----84
Db	6458	NAEVSLECELSGTPPPFVWVWKDKRQLRSSKKYKIASKNPHTSIHILNVDTSIDIGEYHCK 6517
QY	85	-----84
Db	6518	AQNEVGSDTCVCTVKLKEPPRFVSKLSLTVWAGEPAELQASIEGAQPIFVOWLKEKEEV 6577
QY	85	-----84
Db	6578	IRESENIRITFVENVATLOFAKAPANAGRYICQIKNDGGMEENMATLMVLEPAVIVEKA 6637
QY	85	-----84
Db	6638	GPMTVTVGEICTLECKVAGTPELSVEWYKDGKLLTSSQKHKFSYKINLSSRLILSVEROQD 6697
QY	85	-----84
Db	6698	AGTVTFQVQNNVGKSCTAVVDVSDRAVPPSFTRRLLKNTGGVLGASCILECKVAGSPIS 6757
QY	85	-----SVLTV- 89
Db	6758	VAMFHEKTIKISGAKYQTTTFSDNVCTIQLNSLSDSDMGNTCTVAANVAGSDECRAVLTQ 6817
QY	90	-----89
Db	6818	EPFSFVKEPEPLEVLPCKNVTFTSVIRGTPPFKVNWFRGARELVKGRDRCNIYFEDTVAEL 6877
QY	90	-----89
Db	6878	ELFNIDISQGEYTCVYVNNAGQASCTTRLFVKPEAFLKRLSDHSVEPGKSIILESTYT 6937
QY	90	-----89
Db	6938	GTLPISVTKKDGFNITTEKCNIVTTEKTCILEILNSTRKDRAGQYSCIEIENEAGRDVCG 6997
QY	90	-----89
Db	6998	ALYSTLEPPYFVTEPLEAAGVDSVLSQCOVAGTPEITVSWYKGDTKLRTPPEYRTYFT 7057
QY	90	-----89
Db	7058	NNVATLVFNKVNINDSGEYTCKAENSIGTASSKTVFRIQERQLPPSFARQLKDIEQTUGL 7117
QY	90	-----89
Db	7118	PVTLTCLRLNGSAPIQVCWYRDGVLLRRDENLQTSFVDNVATLKLQTLDSHSGOYSCSAS 7177
QY	90	-----89
Db	7178	NPLGTASSARLTAREPKSPFFDIKPYSIDVIAGESADFECHVTGAQPMRITWSKDNKE 7237
QY	90	-----89

Db 7238 IRPGNYTITCVGNTPLHRLKLVKGDGSGYTCQATNDYKDMCSAQLSVKBPFFVKKL 7297  
QY 90 -----LHODW----- 94  
Db 7298 EASKVAKGESIQLECKISGSPKIVSWFRNDSSELHESWKYMSFINSVALLTINEASAE 7357  
QY 95 -----II:----- 94  
Db 7358 DSGDYICEAHNGVGDA SCSTALTVRKPPVFTQKPSVGALKGSDVILQCEISGTPPFV 7417  
QY 95 ----- 94  
Db 7418 WVKRQVRNKKFKITTSKHFDTSHLNLNLEASDVGEYHCKATNEVSGDSCSVKFKEP 7477  
QY 95 ----- 94  
Db 7478 PRFVKKLSDTSLIGDAVELRAIVEGFQPIVVVWLKDRGEVIRESENTRISFDINATIQ 7537  
QY 95 ----- 94  
Db 7538 LGSPEASNSGKYICQIKNDAGMRECSAVLTVLEPARIEKPEPMTVTTGNPFALBCVVTG 7597  
QY 95 ----- 94  
Db 7598 TPESAKWFKDGRSADSCKHITFINKVASLKIPCAEMSDKGLYSFEVKNVSGKSNCTV 7657  
QY 95 ----- 94  
Db 7658 SVHVSDRVPPSFIRKLDVNALGASVLECRVSGSAPISVGFQDQNEIYVSGPKQSS 7717  
QY 95 ----- 94  
Db 7718 FSENVCTNLNLLPESDGTGIYTCVAANVAGSDECSAVLTQBPSPSEQTPDSVEVLPGMS 7777  
QY 95 ----- 94  
Db 7778 LFTSVIRTPPFKWKFGSRELVPGESCNISLEDFVTELELFEVQPLESGDYSCLVTN 7837  
QY 95 ----- 94  
Db 7838 DAGSASCTHFLFKBPATFVKRLADFSVETGSPVILEATYTGTPPISVSWIKDEVILISQS 7897  
QY 95 ----- 94  
Db 7898 ERCSITMTEKSTILELESTIEDYQYSLIENEAGQDICEALVSLEPPYFIEPLEHVE 7957  
QY 95 ----- 94  
Db 7958 AVIGEPATLQCKVDGTPEIRISWYKEHTKLSAPAYKMQFKNNVASLVINKVDHSDVGEY 8017  
QY 95 -----LNGKE----- 99  
Db 8018 SCKADNSVGAVASSAVLVIKERKLPFFFARKLKDVEHTLGPVAFECRINGSLEPQVSWY 8077  
QY 100 -----:--- 99  
Db 8078 KDGVLKDDANLQTSFVHNVAFLQILQTDQSHIGQYNCASNPLGTASSAKLILSEHEV 8137  
QY 100 ----- 99  
Db 8138 PPFOLKPVSDALGESGTFKCHVTGTAPKITWAKDNREIRPGNGYKMTLVENTATLT 8197  
QY 100 ----- 99  
Db 8198 VLKVGKGDAGQYTCYASNAGKDCSAHLGVQEPFRFIKKLEPSRIVKQDEFTRYECKIG 8257  
QY 100 ----- 99  
Db 8258 GSPKIKVLWKDETEIQSSSKFRMSFVDSVAVLEMHNLSVEDSGDYTCEAHNAAGSASS 8317  
QY 100 ----- 99  
Db 8318 TSLKVKPEPIFRKKPHPIETLKGADVHLECELOQTTPPFHVSWKDKRELRSKKYKIMSE 8377

QY 100 ----- 99  
Db 8378 NFLTSIHILNVDAADIGEYQCKATNDVGDTCVGSIALKAPPRFVKKLSDISTVTVGKEVQ 8437  
QY 100 ----- 99  
Db 8438 LQTTIEGAEPISVWFKDKGEIVRESNDNIWISYSENIATLQFSRVEPANAGKYTCQIKND 8497  
QY 100 ----- 99  
Db 8498 AGMQECFATLSVLEPATIVERKESIKVTTGDTCTTLECTVAGTPELSTKWFKDGLKELTSDN 8557  
QY 100 ----- 99  
Db 8558 KYKISFFNKVSGKLIINVAPSDSGVYSPFVQNPVQKDSCTASLQVSDRTVPSPFTRKLKE 8617  
QY 100 ----- 99  
Db 8618 TNGLSGSSVWMECKVYGSPPISVSWFHEGNEISSGRKYQTTLTDTNTCALTVNMLEESDSG 8677  
QY 100 ----- 99  
Db 8678 DYTCTIATNAGSDECSAPLTVREPPSFVQKPDMDVLTGTNTFTSIIVKGTPTPFSVSWFK 8737  
QY 100 ----- 99  
Db 8738 GSSELVPGDRCNVSLSDSVALELFDVDTQSGETYCIVSNKAGKASCTTHLYIKAPAKF 8797  
QY 100 ----- 99  
Db 8798 VKRLNDYSIEKGKPLILEGTFTGTPPISVTWKNGINVTSPORCNITTTKSAILEIPSS 8857  
QY 100 ----- 99  
Db 8858 TVEDAGQYNCYTENASGKDCSAQILILEPPYFVKQLEPVKVSVDASLQCLAGTPEI 8917  
QY 100 ----- 99  
Db 8918 GVSWKGDTKLRPTTYTKMHFNNAVTLVFNQVDINDSGEYICKAENSVEGVSASTFLTV 8977  
QY 100 ----- 99  
Db 8978 QEQKLPPSFRQLRDVQETVGLPVVFDCAISGSEPSVSWYKDGKPLKDSNVQTSFLDN 9037  
QY 100 ----- 99  
Db 9038 TATLNIFKTRSLAGQVSCYCTATNPIGSASSARLILTEGKNPPFFDIHLAPDAVVGESA 9097  
QY 100 ----- 99  
Db 9098 DFECHVTCTQPIKYSWAKDSREIRSGGKYQISYLENSAHLTVLKVDKGDGSGYTCYAVNE 9157  
QY 100 ----- 99  
Db 9158 VGKDSCTAQLNIKERLIPPFTKRLSETVEETEGNSFKLEGRVAGSQPITVAVYKNNIEI 9217  
QY 100 -----YKCKVSN----- 106  
Db 9218 OPTSNCETTFKNNTLVLQVRKAGMDAGLYTCVSNNDAGSALCTSSIVIKEPKKPPVFDQ 9277  
QY 107 ----- 106  
Db 9278 HLTPTVTSGEYVOLSCHVQSEPIRQWLKAGREIKPSDRCSFSPASGTAVLELRVAK 9337  
QY 107 ----- 106  
Db 9338 ADSDGYVKASNVAGSDTTKSKVTIKDKPAVAPATKAAVDGRLFFVYSEPOSIRVVEKTT 9397  
QY 107 ----- 106  
Db 9398 ATFTIAKVGDPFIPNVKWTGKWROLNQGRVFIHQKGDEAKLEIRDTTKTDSGLYRCVAF 9457

QY	107	-----	106
Db	9458	NEHGEIESNVNLDQERKKQKIEGDLRAMLKKTPIILKKGAGEEEDIDIMELLKNVDPK	9517
QY	107	-----	106
Db	9518	YEKARYMIGITDFRGLQAFELLKQSEETHRLIEIETIERSRDEKEFEELYSFIQORL	9577
QY	107	-----	106
Db	9578	SQTEPVLTKIDENQIVLKDNDVAFIDIKINYPEIKLSWKYKTEKLEPSDKFEISIDGD	9637
QY	107	-----	106
Db	9638	RHLRVKNCOLKDOGNVRLVCGPHIASAKLTVIEPAWERHLQDVTLKEGTCMTQCFSV	9697
QY	107	-----	106
Db	9698	PNVKSEFRNGRILKPOGRHKTEVEHKVHKLTIADVRAEDQGYTCKYEDLETSaelRIE	9757
QY	107	-----	106
Db	9758	AEPIQTKRIONIVVSEHQSAFECESVDDAIVTWYKGPTELTESQYNFRNDGRCHYM	9817
QY	107	-----	106
Db	9818	TIHNVTPDDEGVYSIARLEPRGEARSTAElyLTtKEIKLELKPDPIDSRVPIPTMPiR	9877
QY	107	-----	106
Db	9878	AVPEEIPPVAPPiPILLPTPEKKPPKRiEVTKAVKDKAKVYAKPKEMTPREEiV	9937
QY	107	-----	106
Db	9938	KKPPPTTLIPAKAPeLiDVSSKAEEVKIMTiTRKKEVQKEAEVYEkQAVHKEKRVFI	9997
QY	107	-----	106
Db	9998	ESPEEPYDELEVPYTEPFQPYIEPDYEEIKVEAKKEVHEEEDPEEQEYERE	10057
QY	107	-----	106
Db	10058	EGYDEGEEMEAEAYQREVIQVQKEVYEEShERKVPKAVPEKKAPPPKVIKKPVIEKIE	10117
QY	107	-----	106
Db	10118	KTSRRMEEEKVQTKVPEVSKKIVPQKPSKTPVQEEVIEVKVPVHTKMWiSEKMFFA	10177
QY	107	-----	106
Db	10178	SHTEEESVTVPEVQKEiVTEEKIHVAVSKRVEPPPKVPeLPEKPAPEAVPPIPKKVE	10237
QY	107	-----	106
Db	10238	PPAPKVPPEPKVPPEKKPVPVPKKEPAAPKVPPEPKVPPEKIPVPVAKKKEAPPA	10297
QY	107	-----	106
Db	10298	KVPEVQKRVVTEKiTiVTQRESPPPAVPEiPKKKVPEKFPVPRKEEVPppPKVPAL	10357
QY	107	-----	106
Db	10358	PKKVPPEKVAVPVPVAKAPPRAEVSKKTVVVEEKSRFVAEKLSFAVQORVEVTRHEVS	10417
QY	107	-----	106
Db	10418	AESEWSYSEEGVSiSVIREEEREEEAeVTEYEVMEPEPEYVVEEKLiHSKRVEAE	10477
QY	107	-----	113
Db	10478	PAEVTERQEKKiVLKPiPAKIEEPPPAKVPEAPKkiTVPEKKVPAPVPKKEKVPppKVP	10537
QY	114	-----	113
Db	10538	EPKKVPPEKKVPVKVIKMBEPPLPAKVTERHMQiTOEBKVLVAVTtKKEAPPKARVPPEPKR	10597
QY	114	-----	113
Db	10598	AVPEEKVLKLPKREEEPPAKVTEFRKRVVKEEKVSiEAPKRBPQPiKEVtIMEEKERAY	10657
QY	114	-----	113
Db	10658	TLEEAVSVQRESEYEEYEDYKPEFEYEPTEYDQYEEYEREYERYEHEEYiTEPE	10717
QY	114	-----	113
Db	10718	KPIPVKVPPEEPVPTKPAKPAKVLKKAPEKVPVPIPKKLKPPPKVPPEPKKVPPEEK	10777
QY	114	-----	113
Db	10778	IRISITKREKQVTEPAAKVPMKPRKRVABEKVPVRKEVAPVVRPEVPEVKELEPEEVAF	10837
QY	114	-----	113
Db	10838	EEEVTHVEEYLYVEESEEYiHEEEFiTBEENVVPIPVKYPEVRKVPPEKKFPVPVK	10897
QY	114	-----	113
Db	10898	KEAPPAKVPEVPKKPEKVPVLIPKKEKPPAKVPEVPKVPPEKVPVPVKKEAPPA	10957
QY	114	-----	113
Db	10958	KVPEVPKKVPPEKKVPVPAPKKEAPPAKVPEVPKLiLPEKKPTVPVKKVEAPPPKVPK	11017
QY	114	-----	113
Db	11018	KREPVPVPVALPQEEEVLFEEiVPEEEVLPBEEVLPBEEVLPBEEVLPBEEVLPBEEI	11077
QY	114	-----	113
Db	11078	EEVPPEEYVPEEEEFVPEEEVLPVKVKVPVPAPVVPVPEIKKKVTEKKVViPKEEAPPA	11137
QY	114	-----	113
Db	11138	KVPEVPKKVEEKRIiLLPKBEEVLPVETPEBEPiSEBEPPEPPSiEEVEEVPAPRVP	11197
QY	114	-----	113
Db	11198	EVIKKAVPEAPTVPVKKVEAPPAKVSKKiPEEKVPVPVQKKEAPPAKVPEVPKVPPEKKV	11257
QY	114	-----	113
Db	11258	LVPKKEAVPPAKGRTVLEEKVSVAFRQEVVVVKERLELVVEAEVVEiPEEEFHEVEEYF	11317
QY	114	-----	113
Db	11318	EEGEFHEVEEFiKLEQHRVDEEHKRVKHViEVFEAEVEVEVKPAKPGPiSEKii	11377
QY	114	-----	113
Db	11378	PPKKPTKVVPKPEPAKVPEVPKKiVVEEKVRVPEPRVPPTKVDPDLPPKEVVPPEKKV	11437
QY	114	-----	113
Db	11438	PVPPAKKEAPPKVPPEAPKEVVPPEKKVPVPPPKPEVPPTKVPEVPAKAAVPEKKVPEAI	11497
QY	114	-----	113
Db	11498	PPKPESPPPPEVPEAPKEVVPPEKKVPAAPKKKPEVTVKVDEAPKEVVPPEKKVPVPPKKP	11557
QY	114	-----	113
Db	11558	EVPTTKVPEVPKVAPEKKVPEAiLPPKPESPPPPEVPEEPEEVALLEEPPAEVVEEPEPAAP	11617
QY	114	-----	113

Db 11618 PQVTPPKKVPPEKKAPAVAKKPELPPVKVPEVPEKKVPLVPPKPEAPPKVP 11677  
QY 114 ----- 113  
Db 11678 EYPKEVPEKKVAVPKKPEVPPAKVPEVPKPKVLEKPAVPVPERAESPPPEVVEPEEI 11737  
QY 114 -----EK 115  
Db 11738 APEEEIAPEEKPVVAEEEEPEVPPAVPEPKKIPEKKVPVKKPEAPPKPEPEK 11797  
QY 116 TISK----- 119  
Db 11798 VIEPKLKRPPPPPPAPPKEDVKEIFQLKAIPKKKVPKQVPEKVELTPLKVPGGEK 11857  
QY 120 ----- 119  
Db 11858 KVRKLLPERKPEKEEVVLKSVLRPEEEEEPKVPEKKLEKVKKPAVPEPPPKPEVEE 11917  
QY 120 ----- 119  
Db 11918 VPTVKREKKIPEPTKVPEIKPAIPLPAPEPKKPEAEVKTIKPPVPEPTPIAAPVTV 11977  
QY 120 ----- 119  
Db 11978 PVVGKAEAKAPKEAAKPKGPIKGVPKKTPSPIEAERRKLRPGSGEKPPDEAPTYQL 12037  
QY 120 ----- 119  
Db 12038 KAVPLKFVKEIKDIILTESEFVGSSAIFELVSPSTAITTMKDGNSIRESKHFRIADG 12097  
QY 120 -----AKGP- 124  
Db 12098 KDRKLHIIDVQLDAGEYTCVLRGNKEKTSKALVVEELPVRFVKTLEEEVTVVKGQPL 12157  
QY 125 ----- 124  
Db 12158 YLSCELNKERDVVWRKDKIVVEKPGRIVPGVIGLMLRALTIINDADTDAGTYTVTENAN 12217  
QY 125 ----- 124  
Db 12218 NLESCCVKVEVIRDVLVKPIRDQHVPKGTAIFACDIADKTPNIKWFKYDEIPAEPN 12277  
QY 125 ----- 124  
Db 12278 DKTEILRDGNHLYLKIKNAMPEDIAEYAVEIEGKRYPAKLITIGEREVELLKPIEDVTIVE 12337  
QY 125 ----- 124  
Db 12338 KESASFDAEISEADIPGOMKLGELLRPSPTCEIKAEGGKRFLTLRKVKLDQAGEVLYQA 12397  
QY 125 ----- 124  
Db 12398 LNAITAILTVKEIELDFAVPLKDVTPVERQARFECVLTREANVWSKGPDIKSSDKF 12457  
QY 125 ----- 124  
Db 12458 DIIADGKKHILVINDSQFDDGGVYTAIEVEGKTSARLFTVGTGIRLFMSPLEDQTVKEGET 12517  
QY 125 ----- 124  
Db 12518 ATFFVCELSHEKMHVVMFKNDAKLHTSRTVLISSEKTHKLEMKEVTLDDISOIKAQVKEL 12577  
QY 125 ----- 124  
Db 12578 SSTAQLKVEADPYFTVKLHDKTAVEKDEITLKCEVSKDVPVKWPKDGEIIVSPKYSIK 12637  
QY 125 ----- 124  
Db 12638 ADGLRILKIKKADLKDKGEYVDCGTDKTKANVTVEARLIKVEKPLYGVEFVGETAHF 12697  
QY 125 ----- 124  
Db 12698 EIELSEPDVHGOMKLGQPLTASPDCEIIEDGKKHILILHNCQLGTMTEVFSQAANAKSA 12757

QY 125 -----REPO----- 128  
Db 12758 ANLKVKELPLIFITPLSDVKVFEKDEAKFECEVSREPTFRWLKCTQEIITGDDREFELIKD 12817  
QY 129 ----- 128  
Db 12818 GTKHSWIKSAAFEDEAKYMPAEDKHTSGKLIIEGIRLKFUTPLKDVTAKEKESAVFTV 12877  
QY 129 ----- 128  
Db 12878 ELSHDNIIRVKWFKNDORLHTTRTSVSMQDEGKTHSTTFKDLSDIDTSQIRVEAMGMSSEAK 12937  
QY 129 ----- 128  
Db 12938 LTVLEGDYPFTGKLDQYTGVEKDEVILQCEISKADAPVKWFKDGKEIKPSKNAVIAKDGK 12997  
QY 129 ----- 128  
Db 12998 KRMLILKKALKSDIGQYTCDCGTDKTSGKLDIEDREIKLVRPLHSVEVMEETARFETEI 13057  
QY 129 ----- 128  
Db 13058 SEDDIHANWKLKGEALLQTPDCEIKEGKIHSVLHNCRLDQGTGVDFQAAVNVKSAHLR 13117  
QY 129 ----- 128  
Db 13118 VKPRVIGLLRPLKDVTVTAGETATFDCELSYEDIPVHWYLGKKLEPDKVVPVRSRSEKGVH 13177  
QY 129 ----- 128  
Db 13178 TLTLRDVKLEADAGEVOLTAQDFKTHANLFVKPEPVEFTKPLEDQTVBEGATAVLECEVSR 13237  
QY 129 ----- 128  
Db 13238 ENAKVWFKNGTEILSKKYEIVADGRVKLVIHDCPTEDIKTYTCDAKFKTSCNLNV 13297  
QY 129 ----- 128  
Db 13298 PPHVEFLRPLDLQVREKEMARFECELSRENKAKWFKDGAEIKGKKYDIISKAVRIL 13357  
QY 129 ----- 128  
Db 13358 VINKCLLDEAEYSCEVRTARTSGMLTVLEEAFTKNLANIEVSETDTIKLVCVSKPG 13417  
QY 129 ----- 128  
Db 13418 AEVIWYKGDEEIIETGRVEIILTEGRKRILVIONAHLEDAGNVCNRLPSSRTDGVKVVHLEL 13477  
QY 129 ----- 128  
Db 13478 AAEFISKPNLEILEGEKAEPVCSISKESFPVOWKRDDKTLESGBKYDVIADGKKRVLV 13537  
QY 129 ----- 128  
Db 13538 KDATLQDMGYVMVGAARAAHLTVIEKLRIVPLKDRVKEQOEVEVFNCEVNTGAKA 13597  
QY 129 -----VYTL----- 132  
Db 13598 KWRNEEAIKSSKVIILQKDLVYTLRIRDAHLDDQANVNSLTNHRGENVKSAAANLIVE 13657  
QY 133 ----- 132  
Db 13658 EEDLRIVEPLKDIEIEMKKSVTFWCKVNRNLNVTLWTKNGEVPFDNRVSRVVDKYKIML 13717  
QY 133 ----- 132  
Db 13718 TKDCGFPDEGEYIVTAGODKSAVELLIIEAPTEFEVHELDQTVTFDQAVFSCOLSREK 13777  
QY 133 ----- 132  
Db 13778 ANVKWYRNGREIKGKKYKFEKDGSIHRLIINKDRLDDCEYACCGVEDRKSARLFPVEI 13837

QY 133 ----- 132  
Db 13838 PVELIRPPQDILEAPGADVFLAELNKDKVEQVQLRNNMVMVVOGDKHOMSECKIHLQI 13897  
QY 133 ----- 132  
Db 13898 CDIKPRDQGEYRFIAKDEARAKELAAAPKIKTADQDLVVDVGKPLTMVVPYDAYPKAE 13957  
QY 133 ----- 132  
Db 13958 AEWFKENEPLSTKIDTTAEQTSFRILEAKKGDKGRYKIVLQNKHGAEGFINLKVIDVP 14017  
QY 133 ----- 132  
Db 14018 GPNRLEVTETFDGEVSLAWEEPITDGGSKIIGYVVVERRDIKRKTWVLATDRAESCFTV 14077  
QY 133 ----- 132  
Db 14078 TGLQKGGVEYLFRVSAARNVGTGEPVETDNPVEARSKYDVGPPPLNVITIDVNRFGVSLT 14137  
QY 133 ----- 132  
Db 14138 WEPPEYDGAETNYVIELRDKTSIRWDTAMTVRAEDLSATVTDVVEGQEYSFRVRAQNR 14197  
QY 133 ----- 132  
Db 14198 IGVKPSAATPFVKVADPIERPSPVNLTSDDQTQSSVLKWEPLKDGSPILGYIER 14257  
QY 133 ----- 132  
Db 14258 CEEKDNWIRCNMKVLPELTYKVTGLEKGNKYLYRVAENKAGVSDPSEILGPLTADDAF 14317  
QY 133 ----- 132  
Db 14318 VEPTMDLSAPKDGLEIVPNPITILVPSTGYPRPTATWCFGDKVLETGDRVKMKTLSAYA 14377  
QY 133 ----- 132  
Db 14378 ELVISPERSDKGIYTLKLENRVKTSIGEIDVNVIAEPSAPKELKFGDITKDSVHLTWE 14437  
QY 133 ----- 132  
Db 14438 PDDGGSPLTGYVVEKREVSRTWTVMDFVTDLEFTVDPDLVQGEYLFKVCARNKCGPG 14497  
QY 133 ----- 132  
Db 14498 EPAIVDEPNVNSTPATVDPDENVKWRDRTANSIFLTWDPPKNDGGSRKIKGYIVERCPRG 14557  
QY 133 ----- 132  
Db 14558 SDKWVACGEPVAETKMEVTGLEGKWYAYRVKALNROGASKPSRPTETIQAVDTQEAPEI 14617  
QY 133 ----- 132  
Db 14618 FLDVKLLAGLTVKAGTKIELPATVGTGPEPKITWTKADMILKODKRITIEVNPVKSTVTI 14677  
QY 133 ----- 132  
Db 14678 VDSKRSDTGYIIIEAVNVCGRATAVEVNVLDKPGPAAFDITDVTNESCLLTWNPPRDD 14737  
QY 133 ----- 132  
Db 14738 GGSKITNVVERRATDSEVWHKLSSTVKDTNFKATKLIPNKEYLFRVAAENMYGVGEPVQ 14797  
QY 133 ----- 147  
PPSR---DELTKNOVSLT  
||:| :||: ||:  
Db 14798 ASPITAKYQDPGPPPTLEPSDITKDAVTLTWCEPDDGGSPITGYWVERLDPDTDKW 14857  
QY 148 ----- 147  
Db 14858 RCNMPVKDTTYRVKGLTNKKYFRVLAENLAGPGPKSSTEPILLIKDPIDPPWPGKP 14917  
QY 148 ----- 147

Db 14918 TVKDVGKTSVRLNWKTRPEHDGAKIESYVIELMKTGTDEWVRVAEGVPTTQHLLPLGLEG 14977  
QY 148 ----- 147  
Db 14978 QEYSFRVRAVNKAGESESPDVLCREKLYPPSPRVLWINITKNTADLKWTVPKED 15037  
QY 148 ----- 147  
Db 15038 GGSPIITNYIVEKRDVRKQHQTVDTTVKDTKCIWTPLTGSLYVFRVAENAIGQSDYTE 15097  
QY 148 ----- 147  
Db 15098 IEDSVLAKDTFTTGPFPYALAVDVTKRHVDLKWEPKNDGGRPIQRYVIEKKERLGLTRW 15157  
QY 148 ----- 147  
Db 15158 VKAGTAGPCNFRVTDVIEGTEVOFVRAENAGVGHPSPEILSIEDPTSPSPPLD 15217  
QY 148 ----- 147  
Db 15218 LHVTDAGRKHIAIAWKPPKNGGSPILGYHVEMCPVCTEKWVRNSRPKDLKFKVEGV 15277  
QY 148 ----- 147  
Db 15278 VPKYVYLRVAVNAIGVSEISENVVAKDPCKPTIDLETHDIIIVIEGKLSIPVPF 15337  
QY 148 ----- 147  
Db 15338 RAVPVPTVSHKDGKEVKASDRLTMKNDHISAHLEVPKSVRADAGIYITILENKLGSATA 15397  
QY 148 ----- 147  
Db 15398 SINVKVIGLPCKDIKASDITKSSCKLTWEPPEFDGTPILHYVLERREAGRRTYIPVM 15457  
QY 148 ----- 147  
Db 15458 SGENKLSWTVKDLTPNGEYFFRVKAVNKVGGEYIELKNPVAQDPKQPPVDVVEHN 15517  
QY 148 ----- 147  
Db 15518 PTAEMTITWKPPPLYDGGSKIMGYIEKIAKGEERWRCNEHLVPILTYTAKGLEEGKEY 15577  
QY 148 ----- 147  
Db 15578 QFRVRAENAAGISPSRATPPTKAVDPIDAPKVILRTSLEVKRGDEIADASISGSPYPT 15637  
QY 148 ----- 147  
Db 15638 ITWIKDENVIPEEIKKRAAPLVRRRKGEVQEEFVLPLTQRLSIDNSKKGESQLRVRD 15697  
QY 148 ----- 147  
Db 15698 SLRPDHLGYMIKVENDHGIAKAPCTVSVDTPGPPINFPVEDIRKTSVLCKWEPLDDGG 15757  
QY 148 ----- 147  
Db 15758 SEIINYTLEKKDKTKPDSEWIVVTSTLRHCKYSVTKLIEGKEYLFRVRAENRFGPGPCV 15817  
QY 148 ----- 147  
Db 15818 SKPLVAKDPFGPPADPKPIVEDVTSNMLYKWNPKDNGSPILGYWLEKREVNSTHWSR 15877  
QY 148 ----- 147  
Db 15878 VNKSLNALKANVDGLLEGLTYVFRVCAENAAAGPKFSPSPDKPTAHDPIPPGPPIPRV 15937  
QY 148 ----- 147  
Db 15938 TDTSSTTIELEWEPFAGGGEIYGVYFDKQLVGTNEWSRCEKMKIVROYTVKEIREGA 15997  
QY 148 ----- 147

Db 15998 DYKLRVSAVNAAGEGPPGETQPTVTAEPQEPAPVELDVSVKGGIQIMAGKTLRIPAVVTG 16057  
QY 148 ----- 147  
Db 16058 RPVPTKWTKEGELDKDRVVIDNVGTSELIIDKALRKHGRYVITATNSCGSKFAAAR 16117  
QY 148 ----- 147  
Db 16118 VEFVDPGVPVLDLKPVTNRKMLNWSODEDDGGSEITGFIILERKDAKMHWTWRQPIETE 16177  
QY 148 ----- 147  
Db 16178 RSKCDITGLLEGOEYKFRVIAKNKFGCGPPVEIGPILAVDPLGPTSPERLTYTERTKST 16237  
QY 148 ----- 147  
Db 16238 ITLDWKEPRNSGSPIOGYIIIEKRRHDKDPFRVNRRLCTPTTSFLVENLDEHOMYFRVK 16297  
QY 148 ----- 147  
Db 16298 AVNEIGESEPSPLNVVIOQDEVPPTIKLRLSVRGDTIKVKAGEPVHIIPADVTGLPMPKI 16357  
QY 148 ----- 147  
Db 16358 EWSKNETVIEKPTDALQITKEEVSRSSEAKTELSIPKAVREDKGTVTYTASNRLGSVFRNW 16417  
QY 148 ----- 147  
Db 16418 HVEVYDRPSPRNLAVTDIAKESCYLTWDAPLDNGGSEITHYVIDKRDASRKAWEVET 16477  
QY 148 ----- 147  
Db 16478 NTAVERKRYGIWKLIPNGOYEFVRVAVNKYGISDECKSVVIOQDPYRLPGPGPKPKVLAR 16537  
QY 148 ----- 147  
Db 16538 TKGMLVSWTPPLDNGGSPITGYWLEKREEGSPYWSRVSRAPIITKVGLKGVEFNVRLLLE 16597  
QY 148 ----- 147  
Db 16598 GVKYQFRAMAINAAGIGPSESDPEVAGDPIFPFGPPSCPCEVKDKTKSSISLGNKPPAK 16657  
QY 148 ----- 147  
Db 16658 DGGSPIKGYIVEMOEBGTTDKRVNEDPKLITTCWCVPNLKELRYRPRVAVNEAGES 16717  
QY 148 -----CLV----- 150  
Db 16718 EPSDTTGEIPANDIQEEPEVFIIDIGAQCCLVCKAGSQIRIPAVIKGRPTPKSSWEFDGKA 16777  
QY 151 ----- 150  
Db 16778 KKAMKGVHDIPEDAQLETAENSSVIIIECKRSHTGKYSITAKNKAGQKTANCVRKVM 16837  
QY 151 -----KGF----- 153  
Db 16838 VPGPPKDLKVSIDITRSGRLSNKMPDDGGDRIKGVIEKRTIDGKAWTKVNPDCGSTTF 16897  
QY 154 -----YPSD----- 157  
Db 16898 VVPDLLSEQQYFRVRAENRFGIGPPVETIQTARTDPIYPPDPPIKLIKIGLITKNTVHL 16957  
QY 158 ----- 157  
Db 16958 SMKPPKNDGSPVTHYVIVECLAWDTGTKEAWROCNRKRDVEELQFTVEDLVEGGEYEP 17017  
QY 158 ----- 157  
Db 17018 VKAVNAAGVSKPSATVGPCDCQDMPSPSIDLKEEFMEVEBGTNVNIVAKIKGVPEPTLTW 17077  
QY 158 ----- 157  
Db 17078 FKAPPKPDNKEPVLVDTHVNLVDDTCLTIVPOSRRSDTGLTYTITAVNNIGTASKEMR 17137

QY 158 ----- 157  
Db 17138 LNVLGRPPVPPIKPESVSADOMTLSNFPKDDGGSKITNVYIEKREANRKTWVVSSE 17197  
QY 158 ----- 157  
Db 17198 PKECTYTIKLLLEGHEYVPRIMAQNKYGIGELDSEPTARNLFSVPGAPDKPTVSSVTR 17257  
QY 158 --IAVEWE----- 163  
Db 17258 NSMTVNWEEPEYDGGSPVTGYWLEMKDTTSKRWKVRNRPDIKAMTILGVSYKVTGLIEGSD 17317  
QY 164 ----- 163  
Db 17318 YQFRVYAINAAGVGASPLSDPATARDPIAPGPPPKVTDWTKSSADLEWSPPLKDGGS 17377  
QY 164 ----- 163  
Db 17378 KVTGYIIVEYKEBGEKBEWEGKDKVEVGTCLVVTGLKEGAFYKFRVSAVNIAGIGEPGEVT 17437  
QY 164 ----- 163  
Db 17438 DVIEMKRLVSPDLQLDASVRDRIVVHAGGVIIRIIAYVSGKPPPTVTWNNNERTLPQEAT 17497  
QY 164 ----- 163  
Db 17498 IETTAISSMWTKNCORSHQGVYSLAKNEAGERKKTIIIVDVLDPGPPVGTPTFLAHLNLTN 17557  
QY 164 -----SNGQPENNY----- 172  
Db 17558 ESCKLTWFSPEDDGGSPITNYVIEKRESRRRAWTPTYTYVTTRQATVQGLIOGKAYFPRI 17617  
QY 173 ----- 172  
Db 17618 AAENSIGMGPFVETSEALVIREPITVPERPEDLEVKEVTKNVTLTWNPYPKYDGGSEIIN 17677  
QY 173 ----- 172  
Db 17678 YVLESRLIGTEFKHVTNDNLLSRKYTVKLGKEGDTYEVRSVAVNIVGOGKPSFCTKPTT 17737  
QY 173 -----KTT 175  
Db 17738 CKDELAPPTLHLDLFRDKLTIRVGEAFALTGRYSGKPKVSWFKDEADVLEDDRTHIKTT 17797  
QY 176 P----- 176  
Db 17798 PATLALEKIKAKRSDSGKYCVVVENSTGSRKGCQVNVVDRPGPPVGVPSFDEVTKYMV 17857  
QY 177 ----- 176  
Db 17858 ISWKPPLDGGSKITNYIIIEKKEVGKDVMPVTSASAKTCKVKSLLEGKDYIFRIHAEN 17917  
QY 177 ----- 176  
Db 17918 LYGISDPLVSDSMKAKDRFRVPDAPDQPIVTEVTKDSALVTWKNKPHDGGKPTNVILEKR 17977  
QY 177 ----- 176  
Db 17978 ETMSKRWARVTKDPIHPYTKFRVPDLLEGQCYEYFRVSAENEIGIDGPPSPKPVFAKOPI 18037  
QY 177 ----- 176  
Db 18038 AKPSPVNAEALDTCNSVDLTWQPPRHGGSKILGYIVEYQKVGDEEWRANHTPESCP 18097  
QY 177 ----- 176  
Db 18098 ETKYKVTGLRDCQTYKFRVLAVNAAGESDPAHVPEPVLVKORLEPPELLIDANMAREQHI 18157  
QY 177 ----- 176  
Db 18158 KVGDTLRLSAIIKGVPPFKVTWKKEDRDAPTAKRIDVTPVGSKLEIRNAHEDGGIYSUT 18217



QY 177 ----- 176  
Db 18218 VNPAGSKTVSVKVLVDKPGPPDLSEVRKSDCYLTWKEPLDDGGSVITNVVVERRD 18277  
QY 177 ----- 176  
Db 18278 VASAQMSPLSATSKKSHFAKHLNEGNOYLFRVAENQYGRGPFVETPKPIKALDPLHPP 18337  
QY 177 ----- 176  
Db 18338 GPPKDLHHVDVKTEVSLVWKNKPDROGGSPITGYLVEYQEEGTQDWIKFKFTVNLECVVT 18397  
QY 177 ----- 176  
Db 18398 GLOQGYFRVRAENIVGLGLPDTTPIECQEKLVPPSVELDVKLTIEGLVWRAGTTVR 18457  
QY 177 ----- 176  
Db 18458 PAIRGVVPFTAKWTDGSEIKTDEHYTVETDNFSSVLTITKNCLRRDTGTYQITVSNAG 18517  
QY 177 ----- 176  
Db 18518 SKTVAVHLTVLDVPGPTGPIINILDTPEHMTISWQPPKDDGGSPVINYIVEKQDTRKDT 18577  
QY 177 ----- 176  
Db 18578 WGVSSGSSKTKLIPHQLKGCEYFRVRAENKIGVPPDLDSTPTVAKHKFSPPPGKP 18637  
QY 177 ----- 176  
Db 18638 VVTDITENAAVSWTLPKSDGGSPITGYMERREVTGKWVRNKTPIADLKFRTGLYEG 18697  
QY 177 ----- 176  
Db 18698 NYEYFRVAENLAGLKPSPSSDPIKACRPIKPPGPIINPKLKDKSRETADLVWTKPLSD 18757  
QY 177 ----- 176  
Db 18758 GGSFILGYVVECQKPGTAOWNRINKDELIRQCAFVPLGIEGNEYRFRKAANIVGEGEP 18817  
QY 177 ----- 176  
Db 18818 RELAESVIAKDILHPPPEVLDVTCRDVITVRVGQTIIRILARVGRPEPDITWTKGVLY 18877  
QY 177 ----- 176  
Db 18878 REKRVDLIQDLPRVELOIKEAVRADHGKYIISAKNSSGHAQGSAINVLDPRPGQNKLK 18937  
QY 177 ----- 176  
Db 18938 TNVTKENCTISWENPLDNGGSEITNFIVEYRKPQKGSIVASDVTKRLIKANLANNEY 18997  
QY 177 ----- 176  
Db 18998 YFRVCAENKVGQGTIETKPTILAINPIDRPGEPENLHIAADKGKTFVYLKWRRPDYDGS 19057  
QY 177 ----- 176  
Db 19058 PNL5YHVERLKGSDDERVHKGSIKETHYMWDRVCVENQIYEFVRQTKNEGGSQWVTE 19117  
QY 177 ----- 180  
Db 19118 EVVVKEDLPVLDLKL5GLTVKAGDTIRLEAGVRGKPPPEVAWTKDKDATDLTRSPRV 19177  
QY 181 -----SD-----GSFFLYSKLV----- 193  
Db 19178 KIDTRADSSKFSLTAKRSDGGKYVVTATNTAGSFAYATVNVLDKPGPVRLKIVDVSS 19237  
QY 194 ----- 193  
Db 19238 DRCTVCWDDPDDGGCEIQNIYILEKCTKRMWSTYSATVLTGCTVTRLIEGNEYIFRV 19297  
QY 194 -----DK----- 195

Db 19298 RAENKIGTGPPTESKPVIAKTKYDKPGRPDPPEVTKVSKBEIMVVMNPPBYDGGKSITGY 19357  
QY 196 ----- 195  
Db 19358 FLEKKEKHSTRWVPVNKSAIPERRMKVONLLPDHEYOFRVKAENEIGIGEPSLSPRPVVA 19417  
QY 196 ----- 195  
Db 19418 KDPIEPPGPTNFRVVDTTKHSITLGMGKPVYDGGAPLIGYVVMERPKIADASPDEGWR 19477  
QY 196 ----- 195  
Db 19478 CNAQAQLVRKEFTVTSIDENQYEFVCAQNOVGIGRPAELKAIPKEILEPEIDLDA 19537  
QY 196 ----- 195  
Db 19538 SMRKLIVRAGCPIRLFAIVRGPRAPKVTWRKVGIDNVVRKGQVDLVDTHAFLVIPNSTR 19597  
QY 196 ----- 195  
Db 19598 DDSKYSLLTVNPAGEKAFVFNVRVLDTPGPVSDLKVSDVTKTSCHVSWAPPENDGGSQV 19657  
QY 196 ----- 195  
Db 19658 THYIVKREADRKTWSTVTPVEVKTSFHTNLPVGEYIFRVTAVNEYGPGVPTDVPKPV 19717  
QY 196 ----- 195  
Db 19718 LASDPLSEPPPRKLEVTEMTKNSATLAWLPLRDGGAKIDGYITSYREEBPADRWTEY 19777  
QY 196 ----- 195  
Db 19778 SVVKDLSLVTLGKEKKYKFRVAARNAGVSLPREAEGYVEAKEQLLPPKILMEQITI 19837  
QY 196 -----SRWQOQN----- 202  
Db 19838 KAGKLRIEAHVYKGPHTCKWKKEDEVVTSSHLAVHKADSSSILIIKDTRKDSGYIS 19897  
QY 203 ----- 202  
Db 19898 LTAENSSGTDQTKIVVMDAPGPPQPFDDISDIDADACLSWHIPLEDDGGSNTNIVE 19957  
QY 203 ----- 202  
Db 19958 KCDVSRGDWITALASVTKTSRVGKLLIPGQEYIFRVRAENRFGISEPLTSPKVAQPFEG 20017  
QY 203 ----- 202  
Db 20018 VPSEPKNARVTKNKDCIFVANDRPDSGGSPIIIGYLIERKERNSLWVKANDTLVRSTE 20077  
QY 203 ----- 202  
Db 20078 YPCAGLVEGLEYSFRIYALNKGSSPPSKPTEYVTARMPVDPDPGKPEVIDVTKSTVSLIW 20137  
QY 203 ----- 202  
Db 20138 ARPKHDGSKIIGYFVEACKLPGBKWVRCNTAPHQIQEYYTATGLEEKAQYQFRAIART 20197  
QY 203 ----- 202  
Db 20198 AVNISPPSEPDVPTILAENVPPRIDLSVAMKSLITVKAGTNVCLDATVFGKMPPTVSWK 20257  
QY 203 ----- 202  
Db 20258 KDGTLLKPAEGIKMAMORNLCLELFSVNRKDSGDTYITAENSSGSKSATIKLKVLDKPG 20317  
QY 203 -----VFSGSV----- 208  
Db 20318 PPASVKINKMYSDRAMLSWEPPELDDGGSEITNYIVDKRETSRPNWAQVSATVPTSCSVE 20377  
QY 209 ----- 208

Db 20378 KIEGHEYOFRICAENKYGVDVETEPAPAKNPYDPPGRCDDPVISNITKDHMTVSWKP 20437  
QY 209 ----- 208  
Db 20438 PADDGGSPITGYLLEKRETOAVNWKVNRKPIIERTLKATGLOEGTEYEFRTAINKAGP 20497  
QY 209 ----- 208  
Db 20498 GKPSDASKAAYARDPOYPAPPAPFKVYDTRSSVSLSWGKPAYDGGSPICGLYEVKRA 20557  
QY 209 ----- 208  
Db 20558 DSDNWWRCNLQONLOKREFTVGLMEDTQYQFRVAVANKIGYSDPSDVPDKHYPKDILIP 20617  
QY 209 ----- 208  
Db 20618 PEGELDADLRKTLILRAGWTMRLYVPVKGRRPPPKITWSKPNVNLDRIGLDIKSTDEDTF 20677  
QY 209 ----- 208  
Db 20678 LRCENVNKYDACKYILTLENSCGKREYTIIVVKVLDTPGPPVNVTVKEISKOSAYVTWEP 20737  
QY 209 ----- 208  
Db 20738 IIDGGSPIINYVQKRDARKSWSTVTTECSKTSFRVANLEEGKSYFFRVAENEYIGID 20797  
QY 209 ----- 208  
Db 20798 PGETRAVKASQTPGVVDLKVRSVSKSCSIGWKPHSDGSGRIIGYVVDFTLTKENKWO 20857  
QY 209 ----- 208  
Db 20858 RVKMSLSQYSAKDLTEGKEYTFRVSAENENGEGTSEITVVARDVVPDLKGLPDL 20917  
QY 209 ----- 208  
Db 20918 CYLAKENSFRUKIPIKGPAPSVSWKKGEDPLATDTRVSVESSAVNTTLIYVDCOKSDA 20977  
QY 209 ----- 208  
Db 20978 GXYTITLKNVACTKEGTISIKVVGKPGIPTGPIKFEVTAEAMTLKWAPPKDDGGEITN 21037  
QY 209 ----- 214  
MHEALH  
: : :  
Db 21038 YILEKRDVNNKWTASAVQKTTFRVTRLHEGMEYTFRVAENKYGVEGLKSEPIVAR 21097  
QY 215 ----- 214  
NHY  
: : :  
Db 21098 HPFDVDPAPPPNIVDVRHDSVSLTWTDPKKTGGSPITGYHLEFKERNLLWKRAKTPPI 21157  
QY 215 ----- 214  
Db 21158 RMRDFKVTGLTBSGLEVEFRVMAINLAGVKPSLPSEPVVALDPIDPPGKPEVINITRNSV 21217  
QY 215 ----- 214  
Db 21218 TLIMTEPKYDGGHKLTYIVEKRDLPKSWMKANHVNVPECAFTVTDLVEGKYEFIRA 21277  
QY 215 ----- 214  
Db 21278 KNTAGAISAPSESTETIICKDEYEAPTIVLPTIKDGLTIKAGDTIVLNAISILGKPLPK 21337  
QY 215 ----- 214  
Db 21338 SWSKAGDIRPSDITQITSTPSMLTIKYATRKDAGEYITATNPFGTKVEHKVTVL 21397  
QY 215 ----- 214  
Db 21398 DVPGPPEISNVAERKATLTWTPLEDGGSPIKSYILEKRETSRLLTWTVVSEDIQSCR 21457  
QY 215 ----- 217  
NHY  
: : :  
Db 21458 HVATKLIQOGNEYIFRVSAVNHYGKEPVQSEPVKVMVDRFGPPGPEKPEVSNVTKNATV 21517

QY 218 ----- 217  
Db 21518 SWKRPVDDGGSEITGYHVERREKKSRLRVRAIKTPVSDLRCKAVTGLQEGSTYEFVRSAEN 21577  
QY 218 ----- 225  
-TQKSLSLS-  
: : :  
Db 21578 RAGIGPPEASDSVLMKDAAYPPGPPSNPHVTDTTKKSASLAWGRPHVDGGLGLEYVVE 21637  
QY 226 ----- 225  
Db 21638 HOKVGDEAWIKDTTGTALRITQFVVPDLQTKYKYNFRISAINDAGVGEPAVIPDVEIVER 21697  
QY 226 ----- 225  
Db 21698 EMAPDFELDAELRRTLIVVRAGLSIRIFVPIKGRPAPEVTWTKDNINLKNRANIENTESFT 21757  
QY 226 ----- 225  
Db 21758 LLLIPECNRYDTGKFVMTIENPAGKSGFVNVRLDTPGPVNLRLPTDITKDSVTLHMDL 21817  
QY 226 ----- 225  
Db 21818 PLIDGGSRTINIVEKREATRKSYSSTATTKCHKCTYKVTGLSEGCYFFRVAENYEGIG 21877  
QY 226 ----- 225  
Db 21878 EPTETTEPVKASEAPSPDLSNIMDITKSTVSLANPKPKHDGGSKITGVVIEAQRKGSQ 21937  
QY 226 ----- 225  
Db 21938 WTHITTVKLECVVRNLTEGEYTFQVMVNSAGRSAPRESRPVIVKEQTMPELDLRCI 21997  
QY 226 ----- 225  
Db 21998 YOKLVIKAGDNIKVEIPLGRPKPTVTKKGQDILKOTQRVNFETTATSTILNINECVR 22057  
QY 226 ----- 225  
Db 22058 SDSGPYPLTARNIVEGVDVITIQVHDIPGPTGPIKFDEVSDDFTVTSWDPPEGGVP 22117  
QY 226 ----- 225  
Db 22118 ISNVVEMRQDSTTWELATTIVIRTTYKATRLTTGLEYQFRVKAQNRVYGVGITSACI 22177  
QY 226 ----- 225  
Db 22178 VANYPFKVPGPPTQVTAVTKDSMTISWHEPLSDGSGPILGYHVERKERNGILMQTVSK 22237  
QY 226 ----- 225  
Db 22238 ALVPGNIFKSSGLTDGIAVEFRVIAENMAGKSKPSKPSEPMALDPIPPGKPVPLNITR 22297  
QY 226 ----- 225  
Db 22298 HTVTILKWAPEYTGFKITSYIVEKRDLPNGRWLKANFSNILENEFTVSGLTEDAAEYFR 22357  
QY 226 ----- 225  
Db 22358 VIAKNAAGAISSPPSEPSDAITCRDDVEAPKIKVDVKFDVTILKAGEAFRLEADVSGRPP 22417  
QY 226 ----- 225  
Db 22418 PTMEWSKDGKELEGTAKLEIKIADFNSTLVNKNDSRRDSGAYTLTATNPGGFAKIIFNVK 22477  
QY 226 ----- 225  
Db 22478 VLDRPGPPEGLAVTEVTSKCVLSWFPPLDDGGAKIDHYIVQKRETSRLAWTNVASEVQ 22537  
QY 226 ----- 225  
Db 22538 VTKLKVTLLKNGNEYIFRVMVKNYGVGEPELSEPVLAENVYGPDPKPNBEVTTITKDS 22597

QY	226	-----	225
Db	22598	MVWGHDPDGGSEIINIVERRDKAGQRIWKCNKKTLDRLYKVSGLTEGHEYEPRI	23657
QY	226	-----	225
Db	22658	AENAGISAPSPSPFYKACDVTFKPGPGNPRVLDTSRSSISIAMNKPIYDGGSEITGY	22717
QY	226	-----	225
Db	22718	MVEIALPEDEWQIVTPPAGLKATSYITIGLTENQEKIRIYAMNSEGLGEPALVPGTPK	22777
QY	226	-----	225
Db	22778	AEDRMLPETELDADLRKVVITIRACCTLRLFPVPIKGRPAPEVKWARDHGESLDKASTEST	22837
QY	226	-----	225
Db	22838	SSYTLIVGNVRPDSKYILTVENSSGSKSAFVNVRLDTPGPPDLKVEVTKTSVTL	22897
QY	226	-----	225
Db	22898	TWDPPLLDGGSKIKNYIVEKRESTRKAYSTVATNCHTSHKVDOLQEGCSYYFRVLAENE	22957
QY	226	-----PGK-----	228
Db	22958	YGIGLPAETAESKASERPLPPGKITLMDVTRNSVLSWEKPEHDGGSRIILGIYIVEMQTK	23017
QY	229	-----	228
Db	23018	GSDKWATCAIVKVTEATITGLIOGEYSFRVSAQNEKGISDPROLSVPVIAKDVIIPPAF	23077
QY	229	-----	228
Db	23078	KLLFNTFTVLAGEDLKVDVDFPIGRPTPAVTHKDNVPLKQTRVNAESTENNSLLTTIKDA	23137
QY	229	-----	228
Db	23138	CREDVGHVVKVLTNSAGEAETLNVILDKPGPTGPVKMDEVTAADTILSWGPKPYDGG	23197
QY	229	-----	228
Db	23198	SSINNYIVEKRDSTTTWQIVSATVARTTIKACRLKTGCEYQFRIAAENRYGKSTYLNSE	23257
QY	229	-----	228
Db	23258	PTVAQYFVKVPGPPGTPVVTLSRDSMEVQWNEPISDGGSRVIGYHLERKERNILWKL	23317
QY	229	-----	228
Db	23318	NKTPQTKFKTTGLEGEVEFRVSAENIVGIGKPSKSECVYARDPCDPPGRPEAIIV	23377
QY	229	-----	228
Db	23378	TRNSVTLOWKKPYDGGSKITGYIVEKKELPEGRWMKASFNTIIDTFEVTGLVEDHRYE	23437
QY	229	-----	228
Db	23438	FRVIARNAAGVFSEPESTGALTARDEVDPPIRISMDPKYKDTIVVHAGESFKVDADIYK	23497
QY	229	-----	228
Db	23498	PIPTIOWIKGDQELSNTARLEIKSTDFATSLSVKDAVRVDSGNVILAKKNVAGERSVTVN	23557
QY	229	-----	228
Db	23558	VKVLDRPGPPGPPVIVSGVTAEKCTLAWKPPLODGGSDIINIYIVERRETSRLVWTVVDAN	23617
QY	229	-----	228
Db	23618	VQTLSCVKTKLLEGNEYTFRIMAVNKYGVGEPLESEPPVAKNPPVDPAPKAPVTTVTK	23677
QY	229	-----	228

---

Db	23678	DSMIVVWERPASDGGSEILGYVLEKRDKEGIRWTRCHKRLIGELRLRVTGLIENHDYEF	23737
QY	229	-----DWLKAIFYD-----	236
Db	23738	VSAENAGLSEPPSPSAIQACDPIYKPGPPNNPKVIDITRSSVFLSKSPKIYDGGCEIQ	23797
QY	237	-----	236
Db	23798	GYIVEKCDVSGEWTMCCTPPTGINKNTNIEVEKLEKHEYNFRICAINKAGVGEHADVPGP	23857
QY	237	-----	236
Db	23858	IIVEEKLEAPDIDLDELRLKIINIRAGGSLRFLVPIKGRPTPEVKWGVGEIRDAALID	23917
QY	237	-----	236
Db	23918	VTSTFTSLVLDNVNRDYGKYTLTFLNSSGTKSAFVTVRVLDTSPPVNLKVITITKDSV	23977
QY	237	-----	236
Db	23978	SITWEPPLDGGSKIKNYIVEKREATRKSAAVVTNCHKNKSWKIDQLQEGCSYYFRVTA	24037
QY	237	-----KVAE-----	240
Db	24038	NEYGIGLPAQTADPIKVAEVPQPPGKITVDVTRNSVLSWTKPEHDGGSKIIOIVIVEMQ	24097
QY	241	-----	240
Db	24098	AKHSEKSECARVKSLOAVITNLQTGEEYLFRRVVAVNEKGRSDPRSLAVPIVAKDLVIEP	24157
QY	241	-----	240
Db	24158	DVKPAPSSYVQVQODLKIEVPISGRPKPITITWKDGLPLKQTRINVTDSLDTLTSIK	24217
QY	241	-----	240
Db	24218	ETHKDDGGYGITVANVVQKTASIEIWTLDKPPKPKVFDVSAESITLSWNPPLYT	24277
QY	241	-----	240
Db	24278	GGCOITNYIVOKRDTTWTVDVWSATVARTLKVTKLKTGTEYQFRIFAENRYGOSFALE	24337
QY	241	-----	240
Db	24338	SDPIVAQYPYKEPGPPGTPFATAISKDSMWIQWHEPVNNGGSPVIGYHLERKERNILWT	24397
QY	241	-----	240
Db	24398	KVNKTIIDHTQFKAQNLKEGIEYEFRRVAENIVGVGKASKNSECYVARDPCDPPGTPEPI	24457
QY	241	-----	240
Db	24458	MVKRNEITLOWTRPVYDGGSMITGYIVEKRDLPDGRWMKASFNTVITQFTVSGLTEDQR	24517
QY	241	-----	240
Db	24518	YEFVIAKNAAGAIKSPSDSTGPTAKDEVELPRISMDPKFRDITVNVNAGETFRLEADVH	24577
QY	241	-----	240
Db	24578	GKPLPTIWLGRDKIEESARCEIKNTDFKALLIVKDAIRIDGGYILRASNVAGSKSFP	24637
QY	241	-----	240
Db	24638	VNVKVLDRPGPPGPPVQVGTSEKSLTWSPPLODGGSDISHYVVEKRETSRLAWTVVA	24697
QY	241	-----	240
Db	24698	SEVVTNSLKVTKLLEGNEYFRIMAVNKYGVGEPLESAPVLMKNPFLVPGPPKSLEVTNI	24757
QY	241	-----	240

Db	24758	AKDSMTVCWNRPSDGGSEIIIGYIVEKRRDRSGIRWIKCNKRRTDLRLRVTLGLTEHVE	24817
QY	241	-----	240
Db	24818	FRVSAENAGVGEPSPATVYKACDPVKPGPTNAHIVDTTKNSITLAWGKPIYDGGSE	24877
QY	241	-----	240
Db	24878	ILGYVVEICKADEEWOIIVTPTQGLRVTRFEISKLEHQBKIRVCALKNKVLGEATSWP	24937
QY	241	-----	240
Db	24938	GTVPKEDKLEAPELDLDSELKGIIVRAGGSARIHIPKGRPTPEITWSREGEFTDKVQ	24997
QY	241	-----	240
Db	24998	IEKGVNYTOLSDNCDNRNDAGKYLKLENSGSKSAFVTVKVLDPGPPONLAVREKRD	25057
QY	241	-----	240
Db	25058	SAFLVWEPPIIDGAKVKNYVIDKRESTRKAVANVSSKSKTSFKVENLTEGAIYFRVM	25117
QY	241	-----	240
Db	25118	AENERGVGVVETVDAVKAAPSPPGKVTLFDVSQTSASLWMEKPEHDGSRVLGVVE	25177
QY	241	-----	240
Db	25178	MOPKGTKEKSIVAESKVCNAVVTGLSSGOEQYQFRVKAENEKSDPRVLGVVPVIAKDLTI	25237
QY	241	-----	240
Db	25238	QPSLKLPTNYSTIOAGEDLKEIPVIGRPNISWVKDGEPLKQTRVNVEETATSTVLH	25297
QY	241	-----	240
Db	25298	IREGNKDDFGKYTVTATNSAGTATENLSVIVLEKPGPPVGVRFDEVSADFVVISWEPPA	25357
QY	241	-----	240
Db	25358	YTGCGQISNYIVEKRDTTTTTHWMSAVVARTTIKIKLKTGTEYQFRIFAENRYGKSAP	25417
QY	241	-----	240
Db	25418	LDSKAVIQVYPFKEPGPGTPEVTSISKQMLVQWHEPVDNGGTKIIGYHLEQKEKNSLT	25477
QY	241	-----	240
Db	25478	WVKLNKTPIDQTKFKTTGLDEGLEEYFKVSAENIVGIGKPSKVSCEFVARPCDPPGRPE	25537
QY	241	-----	240
Db	25538	AIVITRNVTLWKKPAYDGGSKIYGYIVEKKDLPDGRMKASFNTVLETFVTSGLVED	25597
QY	241	-----	240
Db	25598	QRYEFRVIARNAAGFSEPSDSSGALTARDEIDAPNASLDPKYKDVIVVHAGETFEVLEAD	25657
QY	241	-----	240
Db	25658	IRGKPIPDVYWSKDGKELEETAARMEIKSTIQKTTLVVKDCIRTDGGQYILKLSNVGGTK	25717
QY	241	-----	240
Db	25718	SIPITVKVLDPRGPPGEPGLKVTGVTAEKCYLAWNPPLODGGANISHYIIEKRETSRLSWT	25777
QY	241	-----	240
Db	25778	QVSTEVALNYKVTKLLPGNEVIFRVMVAVNYGIGEPLESGPVTACNPYKPPGPSTPEV	25837
QY	241	-----	240
Db	25838	SAITKDSMVVTWARVPVDDGGTEIEGYILEKRDKEGVRWTKCNKKTLDLRLRVTLGLTEGH	25897
QY	241	-----	240
Db	25898	SYEFRVAAENAGVGEPSPSVYRACDALYPPGPPSNPKVTDTSRSSVSLAWSKPIYDG	25957
QY	241	-----	244
Db	25958	GAPVKGYVVEVKEAAAADEWTTCTPTGLQGQKFTVTKLKENTENYFRICAINSEGVGEPA	26017
QY	245	-----	244
Db	26018	TLPGSVVAQERIEPPEIELDADLRKVVLRSATLRLFVTIKGRPEPEVKEKAEGLTD	26077
QY	245	-----	244
Db	26078	RAQIEVTSFTMLVIDNVTRDPSGRYNLLENNSGSKTAFVNVRLDPSAPVNLTIREV	26137
QY	245	-----	244
Db	26138	KKDSVTLISWEPLIDGAKITNYIVEKRETRKAYATITNNCTKTTFRIENLQEGCSYVF	26197
QY	245	-----	244
Db	26198	RVLASNEYGIGLPAETTEPVPKVSPPPLPPGRVTLVDVTRNTATIKWKPESDGGSKITGY	26257
QY	245	-----	244
Db	26258	VVEMOTKGSEKWTCTQVKTLEATISGLTAGEEYVFRVAANVEKGRSDPRQLGVPVIARD	26317
QY	245	-----	244
Db	26318	IEIKPSVELPHTFNVKAREQLKIDVPFKRPOATVNNRKGQTLKETTRVNVSSKVT	26377
QY	245	-----	244
Db	26378	SLSIKEASKEDVGYELCVSNSAGSITVPTIIVLDRPGPPGPIRIDEVSCSITISNWP	26437
QY	245	-----	244
Db	26438	PEYDGGCOISNYIVEKKETTSTTHIVSOAVARTSIKIVRLTTGSEYQFRVCAENRYGKS	26497
QY	245	-----	244
Db	26498	SYESSAVVAEYFPFPPGPGTPKVVHATKSTMLVTQVPVNDGGSRVIGYHLEYKERSS	26557
QY	245	-----	244
Db	26558	ILWSKANKILITADTQMKVSGLDGLMYEYRVYAENIAGIGKCSKCEPVPARDPCDPPGQ	26617
QY	245	-----	244
Db	26618	PEVTNITRKSYSKWSKPHYDGGAKITGYIVERRELPGRWLKCNTYNIQETTYFEVTELT	26677
QY	245	-----	244
Db	26678	EDQRYEFRVFAARNAADSVSESESTGPIIVKDDVEPPRVMDVKFROVIVVKAGEVLKIN	26737
QY	245	-----	244
Db	26738	ADIAGRPLFVISWAKDGIIEERARTEIISTDNHTLLTVKDCIRRDGTQYVLTLLKNVAGT	26797
QY	245	-----	244
Db	26798	RSVAVNCKVLDKPGPPAGPLEINGLTAECSLSWGRPOEDGGADIDYIIVEKRETSHLAW	26857
QY	245	-----	244
Db	26858	TICEGELQWTSCKVTLLKGNEYIFRVTGVNKYGVGEPLSVAIKALDPFTVPSPPTSLE	26917
QY	245	-----	244
Db	26918	ITSVTKESMTLWCSRPESDGGSEISGYIIERREKNSLRWVRNKKPVYDLRVKSTGLREG	26977

QY	245	-----	244
Db	26978	CEYERYVAENAGLSLPSLRAEDPVFLPSPSPSKIVDSGKTTITIAWVKPLFD	27037
QY	245	-----	244
Db	27038	GGAPITGYTVEYKSSDDTKTSIQSLRGTEYITISGLTGAEYVFRVKSYNKVGASDPDS	27097
QY	245	-----	244
Db	27098	SSDQIAKEREBEPLFDIDSEMRKTLIVKAGASFTMTVPRGRPVNVLSKPEDTDLRTR	27157
QY	245	-----	244
Db	27158	AYVDTTDSRTSLTIENANRNSGKYTLTIQNVLSAASLTILVVKVLDTPGPPTNITVQDVT	27217
QY	245	-----	244
Db	27218	KESAVILSWDVPENDGGAPVKNYHIEKREASKAWSVTNNCRLSKYVTNLQEGAIYFR	27277
QY	245	-----	244
Db	27278	VSGENERGVCIPAETKEGVKITEKPSPEKLGVTISKDSVSLTWLKPEDHGGSRIVHYV	27337
QY	245	-----	244
Db	27338	VEALEKGQKNWKAVALKSTHHVVSGLRSENYFFRFAENQAGLSDPRELLLPVLIKEQ	27397
QY	245	-----	244
Db	27398	LEPPEIDMKNFSPHTVYVRAGSNLKVDIPISGRPLPKVTLSDRGVPLKATMRFNTEITAE	27457
QY	245	-----	244
Db	27458	NLTINKESVTADAGREYTAANSSGTTKAFINIVLDRPGPTGPVVISDITEESVTLK	27517
QY	245	-----	244
Db	27518	WEPPKYDGSQVTNYILLKRETSTAVWTEVSATVARTMMKVMKLTGEEYQFRIKAENRF	27577
QY	245	-----	244
Db	27578	GISHIDSACVTVKLPYTPPGPSTPWNTVNTRESITVGNHPEVNSGSAVGVHLEMKD	27637
QY	245	-----	244
Db	27638	RNSILWQANKLVIRTTHEKVTITISAGLIYEFVYAENAGVGKPSHPSEPVLAIDACEP	27697
QY	245	-----AFD- 	247
Db	27698	PRNVRTIDISKNSVLSWOOPAFDGGSKIRGYIVERRDLPDGRWTKASFNTVTETQFIIS	27757
QY	248	-----	247
Db	27758	GLTONSOYERFVRNAGVSGISNPSEVGPITCIDSYGGPVIDLPLEYTEVWKYRAGTSV	27817
QY	248	-----	247
Db	27818	KLRAGISGKPAPTIEWYKDDKELQTNALVCVENTDLDASILIKDADRLNSGCYELKLRNA	27877
QY	248	-----	247
Db	27878	MGSASATIRVQILDKPGPGGPIEFKVTVAEKITLLWRPPADGGAKITHYIVEKRETSR	27937
QY	248	-----	247
Db	27938	VVMSWSEHLEECIITFTKIIGNEYIFRVRVAVNKYIGEPLESDSVVAKNAFVTPGPPG	27997
QY	248	-----	247
Db	27998	IPEVTKITKNSMTVVWSRPIDAGGSDISGYFLEKRDKKSILGWPKVLKETIRDRQKVTGL	28057
QY	248	-----	247

Db	28058	TENSQYRYCAVNAAGQPFSEPFYKAADPIDPPGPAKIRIADSTKSSITLGWSKP	28117
QY	248	-----	247
Db	28118	VYDGSASVTGYVVEIROGEEETVSTKGEVTRTEYVVVNLKPGVNYFRVSAVNCAGQ	28177
QY	248	-----	247
Db	28178	GEPIEMNEPVOAKDILEAPEILDVALRTSVIAKAGEDVQVLIPFKGRPPPTVTRKDEK	28237
QY	248	-----	247
Db	28238	NLGSDARYSTENTDSSLLTIPOVTRNDTKYILTIENGVEPKSSTSVKVLDTPAACQ	28297
QY	248	-----	247
Db	28298	KLOVKHVSRTVTLLMDPPLIDGGSPIINYVIEKRDATKRTWSVSHKCSSTSKLIDL	28357
QY	248	-----	247
Db	28358	EKTPFFRVLAEINEIGEGEPCETTEPVKAAEVPAPIRDLMSKDKTSTVILSWTKPDPDG	28417
QY	248	-----	247
Db	28418	GSVITEYVVERKKGQETWSHAGISKTCIEIEVSQLEQSVLEFRVFAKNEKGLSDPVTIG	28477
QY	248	-----	247
Db	28478	PITVKELIITPEVDLSDIPGAQVTVRIGHNVHLELYKGRKPSISWLKDLPLKSEFV	28537
QY	248	-----	247
Db	28538	RFSKTENKIITLSIKNAKKEHGGKYTVILDNAVCRIAVPIVITLGPSPKPGPIRDEIK	28597
QY	248	-----	247
Db	28598	ADSVILSWDVPEDNGGGEITCYISIEKRETSQTNWKMVCSSVARTTFKVPNLVKAQYQFR	28657
QY	248	-----	247
Db	28658	VRAENRYGVSOPLVSSIIIVAKHOFRIQPGPKPVYIYNTSDGMSLTWDAPVYDGGSEVTG	28717
QY	248	-----	247
Db	28718	FHVEKKERNILMQKVNTSPISGREYRATGLIVEGLDYQFRVYAENSAGLSSPSDPSKFTL	28777
QY	248	-----	247
Db	28778	AVSPVDPPTPDYIDVTRTITLKNPPLRDGSGKIYGVYSIEKROGNERWRCNFTDVSE	28837
QY	248	-----	247
Db	28838	CQYTVTGLSPDRYEFRIIARNVAGTISPPSSQSGIIMTRDENVPPIVEFGPEYFDGLII	28897
QY	248	-----	247
Db	28898	KSGESLRKALVQGRPVPRVTFWKDGVIEIKRMNMEITDVLGSTSLFVRDATTRHGRVYT	28957
QY	248	-----	247
Db	28958	VEAKNASGAKAEIKVKVQDTPGKVVGPIRFTNITGEKMTLWMDAPLNDGCAPITHYIIE	29017
QY	248	-----	247
Db	29018	KRETSRLAWALIEDKCEAQSYTALKLINGNEYQFRVSAVNVKFGVGRPLDSDPVVAQIQYT	29077
QY	248	-----	247
Db	29078	VPDAPGIPESNITGNSITLTWARPESDGGSEIOQYILERREKKSTRWVKVSKRPISET	29137
QY	248	-----	247

Db 29138 REFVTCLEGEYEFHVMAENAGVGPASGISRLIKCREPVNPPGPTVVKVTDTSKTTV 29197  
QY 248 ----- 247  
Db 29198 SLEWSPKPFVDCGMEIIGYIEMCKADLGDWHKVNAAECVKTRYTVDLQAGEYKFRVSA 29257  
QY 248 ----- 247  
Db 29258 INGAGKDSCEVTGTIKAVDRLTAPELDIDANFKOTHVVRAGASIRLFAYQGRPTTAV 29317  
QY 248 ----- 247  
Db 29318 WSKPDSNLSLRADIHTDSFSTLVENCNRNDAGKYTLTVENNSSKSIITFTVKVLDTPG 29377  
QY 248 ----- 247  
Db 29378 PPGPITFKDVRGTSATLMDAPLLDGGARIHHYVVEKREASRSQWVISEKTRQIFKVN 29437  
QY 248 ----- 247  
Db 29438 DLAEGVYYFRVSAVNEYGVEPEYEMPEIVATEQPAPPRLLDVDTSKSSAVLAWLKPD 29497  
QY 248 ----- 247  
Db 29498 HDGSRITGYLLEMRQKGSDFWEAGHTKQLTFTVRLVEKTEYEFVRKAKNDAGYSEPR 29557  
QY 248 ----- 247  
Db 29558 EAFSSVIKEPOIEPTADLTGITNOLITCKAGSPFTIDVPISGRPAKPVTKLEEMRLKE 29617  
QY 248 ----- 247  
Db 29618 TDRVSITTTKDRTTLTVKDSMRGDSGRYFLTLENTAGVTFSTVWVIGRPGVTGPIEV 29677  
QY 248 ----- 247  
Db 29678 SSVSAESCVLSWGEKPKDGGTEITNYIVEKRESGTTANOLVNSSVKRTQIKVTHLTKYME 29737  
QY 248 ----- 247  
Db 29738 YSFRYSSENREGVKPLESAPIIAEHFPVPSAPTRPEYHVHSANAMSIRWEEPHDGS 29797  
QY 248 ----- 247  
Db 29798 KIIGYWEKKEKERTILWVKENKVPCLCNYKVTGLVEGLEQYFRFYALNAAAGSKASEAS 29857  
QY 248 ----- 247  
Db 29858 RPIMAQNVPDAGRPEVTDVTRSTVSLIWSAPAYDGGSKVYGYIIERKPVSEVGDGRWLK 29917  
QY 248 ----- 247  
Db 29918 CNYTIVSDFVTALSEGDTYEFVRVLAKNAGVISKSESTGPVTCRDEYAPPKAELDA 29977  
QY 248 ----- 247  
Db 29978 RLHGLDVTIRAGSLDLDAVGGKPEPKIITWTKGDKELDCEKVSLOYTKGRATAVIFKC 30037  
QY 248 ----- 247  
Db 30038 DRSDSGKYLTVKNASGTVKAVSMVKVLDSPGCGKLTVSRTQEKTLAWSLPQEDGGA 30097  
QY 248 ----- 247  
Db 30098 EITHYIVERRETSLRNWVVEGECTLSYVVYTRLIKNEYIFRVRANKYGPVGPVESEP 30157  
QY 248 ----- 247  
Db 30158 IVARNSFTIPSPGPPEVGTGKEHIIQWTKPESDGGNEISNLYLDKREKKSRLWRVN 30217  
QY 248 ----- 247  
Db 30218 KDYVVYDTRLKVTSLMGCDYQFRVTAVNAAGNSEPSEASNFISCREPSYTPGPPSAPRV 30277

QY 248 -----WLKAFYD----- 254  
Db 30278 VDTTKHSISLAWTKPMYDGGDIVGYVLEMOEKDQDQWVRVHTNATIRNTEFTVPLDKMG 30337  
QY 255 ----- 254  
Db 30338 OKYSFRVAANVKMGSEYSESAIEIEPVERIEIPDLELADLLKKTVTIRAGASRLMVSFV 30397  
QY 255 ----- 254  
Db 30398 SGRPPPVITWSKQIDILASRAIIDTTESYSLLIYDKVNRDAGKYTIEAENOSGKKSATV 30457  
QY 255 ----- 254  
Db 30458 LVKYVDTPGCPSPVKVKEVSRDSVTITWEIPTIDGGAPVNNYIVEKREAAAMRAFKVTITK 30517  
QY 255 ----- 254  
Db 30518 CSKTLIRISGLVEGTMYYFVRLPENIYGIGPCETSDAIVLSEVPLPAKLEVVDTKST 30577  
QY 255 ----- 254  
Db 30578 VTLAWEKPLYDGGSLTGYVLEACKAGTERMKVVTLPKTVLEHTVTSLSNEGEQYLFRR 30637  
QY 255 ----- 254  
Db 30638 AQNEKGVSEPRETVATVQDLRVLPITDIDTSTMPQKTIHVPAGRPVELVPIAGRPPAA 30697  
QY 255 ----- 254  
Db 30698 SWFFAGSKLRESERVVETHTKVAKLTIRETTIRDTGEYTLLELKNVTGTTSETIKVIIL 30757  
QY 255 ----- 254  
Db 30758 KPGPTGPIKIDEATSITISWEPPELDGGAPLSGYVVEQDRAHRPGWLVPVSEVTRST 30817  
QY 255 ----- 254  
Db 30818 FKFTLREGNEYVFRVAATNRFIGISYLOSEVIECRSSIRIPGPPETLIQIFDVSRDGNTL 30877  
QY 255 ----- 254  
Db 30878 TWYPPEDDGGSQVTGYIVERKEVRADRWVRVNVKVPVTMYRSTGLTEGLEVEHVRVTAIN 30937  
QY 255 ----- 254  
Db 30938 ARGSGKPSRPSKPIVAMDPIAPPQKPNPRVTDTRTSVSLAWSVPEDEGSKVTGYLIE 30997  
QY 255 -----KVAEK 259  
Db 30998 MQYDQHEWTKCNTPTTKIREYTLTHLPQGAERYFRVLACNAGGPEAEVPGTVKVTM 31057  
QY 260 L----- 260  
Db 31058 LEYPDYELDERYQEGIFVROGGVIRLTPIKGPFPICKWTKEGQDISKRAMIATSETHT 31117  
QY 261 ----- 260  
Db 31118 ELVITKEADRGDSGYDLVLENKCKKAVYIKVRVIGSPNSPEGPLEYDDIQVRSVRVSWR 31177  
QY 261 ----- 260  
Db 31178 PPADGGADILGYILERREVPKAAWYITDSRVGTSLVVKGLENVEHFRVSAENQFI 31237  
QY 261 ----- 260  
Db 31238 SKPLKSEEPVTPKPLNPPPPSNPPEVLDVTKSSVLSWSRPPKDDGGSRTVGYIERKE 31297  
QY 261 ----- 260  
Db 31298 TSTDKNVRHNKTOITTTMYTVTGLVPAEYQFRIIAQNDVGLSETSPASEPVCKDPDK 31357

QY 261 ----- 260  
Db 31358 PSQGELEILSISKDSVTLOWEPCDGGKEILGYWEYRQSGDSAWKSKNERIKDKQF 31417  
QY 261 ----- 260  
Db 31418 TIGGLEATEYEFVFAENETGLSRPRTAMSIKTLTSGEAPGIRKEMKDVTTKLGEAA 31477  
QY 261 ----- 260  
Db 31478 QLSQIVGRPLDIIKWRFGKELIQSRKYMSSDGRTHLTVMTEQEDGEGVYTCIATNE 31537  
QY 261 ----- 260  
Db 31538 VGEVETSKLLQATPOFGYPLKEKYGAVGSTLRHVMYIGRPVPMATWFGKLLQ 31597  
QY 261 ----- 260  
Db 31598 NSENITIENTEHTYHLVMKNVQRTHAGKYKVOLSNVFGTVDAILDVEIQDKPKPTGPI 31657  
QY 261 ----- 260  
Db 31658 VIEALLKNSAVISWKKPPADGGSWITNYVVEKCEAKEGAEWQLVSSAISVTTCRIVNLTE 31717  
QY 261 ----- 260  
Db 31718 NAGYFPRVSAQNTFGISDPLEVSSVLIKSPFEKPGAPKPTITAVTKDSCVVAVKPPAS 31777  
QY 261 ----- 260  
Db 31778 DGGAKIRNYYLEKREKKONKWIWTTBEIRETVFSVKNLIEGLEFVRKCNLGGSEW 31837  
QY 261 ----- 260  
Db 31838 SEISEPITPKSDVPDQAPHEKEELNLRVYQSNATLVCKVTHGHPKPIVWYQKREIIA 31897  
QY 261 ----- 260  
Db 31898 DGLYRIQFKGQVHOLIISVTDATVYQVRATNNGSGSVGTASLEVPAPAKIHLPKT 31957  
QY 261 ----- 260  
Db 31958 LEGMAVHALRGEVWSIKIPFGKPDVITWQKQDLIDNNGHYQVIVTRFTSLVFPNG 32017  
QY 261 ----- 260  
Db 32018 VERKAGFYVVCANREGIDQKTVELDVADVPDPPRGVKYSDVSRDSVNLTWTEPASDGG 32077  
QY 261 ----- 260  
Db 32078 SKITNYIVEKCATTAERWLRVQARETRYVINLFGKTSYQFRVIAENKFGLSKPSPE 32137  
QY 261 ----- 260  
Db 32138 PTITKEDKTRAMNYDEEVDRETVSMTKASHSTKELYEKYIMAEIDLGRGFGIVHRCVE 32197  
QY 261 ----- 260  
Db 32198 TSSKITYMAKRVKGTQDQVLVKKEISILNIARHNLHLHESFESMEELVMEFISGL 32257  
QY 261 ----- 260  
Db 32258 DIFERINTSAFELNREIVSVHVCALQFLSHNIGHFDIRPENIYQTRRSSTIKII 32317  
QY 261 ----- 260  
Db 32318 EFGOARQLKPDNRLFTTAPYYAPEVHQHDVVSTATDMWSLGLTVYLLSGINPFLEA 32377  
QY 261 ----- KEAF 264  
: : :  
Db 32378 TNOQIENIMNAEYTFDEEAF 32398

RESULT 14  
010466 PRELIMINARY; PRT: 26926 AA.  
ID Q10466;  
AC Q10466;  
DT 01-NOV-1996 (TrEMBLrel. 01, Created)  
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)  
DE 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)  
DE Titin, heart isoform N2-B (EC 2.7.1.-) (Connectin).  
OS Homo sapiens (human).  
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
OX NCBI\_TaxID=9606;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC TISSUE=HEART;  
RX MEDLINE=96026330; PubMed=7569978;  
RA Labeit S., Kolmer B.;  
RT "Titins: giant proteins in charge of muscle ultrastructure and  
RT elasticity";  
RL Science 270:293-296(1995).  
RN [2]  
RP SEQUENCE OF 2277-25376 FROM N.A.  
RX MEDLINE=92258380; PubMed=1582406;  
RA Labeit S., Gautel M., Lakey A., Trinick J.;  
RT "Towards a molecular understanding of titin.";  
RL EMBO J. 11:1711-1716(1992).  
RN [3]  
RP SEQUENCE OF 1976-2014 FROM N.A.  
RA Labeit S.;  
RL Submitted (DEC-1994) to the EMBL/GenBank/DBJ databases.  
RN [4]  
RP CHARACTERIZATION.  
RX MEDLINE=95331314; PubMed=7607248;  
RA Gautel M., Castiglione-Morelli M.A., Pfuhl M., Motta A., Pastore A.;  
RT "A calmodulin-binding sequence in the C-terminus of human cardiac  
RT titin kinase";  
RL Eur. J. Biochem. 230:752-759(1995).  
CC -1- FUNCTION: THIS GIANT MUSCLE PROTEIN MAY BE INVOLVED IN MUSCLE  
CC ASSEMBLY AND IN MAINTAINING THE STRUCTURAL INTEGRITY OF  
CC SARCOMERES. MAY HAVE PROTEIN KINASE ACTIVITY.  
CC -1- ALTERNATIVE PRODUCTS: A NUMBER OF FORMS OF THIS PROTEIN ARE  
CC PRODUCED BY ALTERNATIVE SPLICING WHICH DIFFER IN TISSUE  
CC DISTRIBUTION. DIFFERENT SIZE TRANSCRIPTS MAY ALSO EXIST WITHIN ANY  
CC ONE TISSUE. THE SEQUENCE SHOWN HERE IS THAT OF THE HEART ISOFORM  
CC N2-B.  
CC -1- TISSUE SPECIFICITY: MUSCLE-SPECIFIC.  
CC -1- SIMILARITY: TO THE CATALYTIC DOMAINS OF OTHER SERINE/THREONINE  
CC KINASES.  
CC -1- SIMILARITY: BELONGS TO IMMUNOGLOBULIN SUPERFAMILY. CONTAINS 112  
CC IMMUNOGLOBULIN C2-LIKE DOMAINS AND 132 FIBRONECTIN TYPE III-LIKE  
CC DOMAINS.  
DR EMBL; X64698; CAA45939.1; -;  
DR EMBL; X83270; CAA58243.1; -;  
DR EMBL; X64697; CAA45938.1; -;  
DR EMBL; X90568; CAA62188.1; -;  
DR EMBL; X64699; CAA45940.1; -;  
DR HSP; P56276; ITLK  
DR InterPro; IPR00282; Cytok\_receptor\_2.  
DR InterPro; IPR000719; Euk\_kinase.  
DR InterPro; IPR000577; FG\_Y\_kin.  
DR InterPro; IPR003962; FNIII\_repeat.  
DR InterPro; IPR003961; FN\_III.  
DR InterPro; IPR001092; HLH\_basic.  
DR InterPro; IPR003598; Ig\_C2.  
DR InterPro; IPR003600; Ig\_like.  
DR InterPro; IPR003006; Ig\_MHC.  
DR InterPro; IPR000129; Peptidase\_S24.  
DR InterPro; IPR002016; Peroxidase.  
DR InterPro; IPR002290; Ser\_thr\_kinase.  
DR InterPro; IPR001245; Tyr\_kinase.  
DR Pfam; PF00041; fn3; 132.  
DR Pfam; PF00047; ig; 91.  
DR Pfam; PF00069; pkinase; 1.





Qy 39 -----EUTC----- 42  
Db 1718 YGVAYSRDSGIITCRATNKYGTDHTSATLIVKDEKSLVEESQLPBGRKGLQRIEELERMA 1777  
Qy 43 ----- 42  
Db 1778 HEGALTGVTTDQKEKQKPDIVLYPEPVRVLEGETARFCRVGTGYPQPKVNWYLNQGLIRK 1837  
Qy 43 ----- 42  
Db 1838 SKRFRVRVDGIHLDIVDCKSYDTGCVKVTAEENPEGVIEHKVKLEIQOREDFRSVLRRAP 1897  
Qy 43 ----- 42  
Db 1898 EPRPEFHVHEPGKLQFEVQKVDPRVDTTETKEVVKLKAERITHEKVPSESELRSKFR 1957  
Qy 43 ----- 42  
Db 1958 RTEGYEAITAVELSKRKKDESYELLRTKDKELLHWTKELEEKKALAEKGITIPT 2017  
Qy 43 ----- 42  
Db 2018 FKPKIELSPSEAPKIFERIQSQTVGQSDAHEFRVRVVGKPDPECEWYKNGVKIERSDR 2077  
Qy 43 -----VVVDVSHED----- 51  
Db 2078 IYWPEDNVCELVIRODTAEDSASIMVKAINIAGETSSHAFLVQAQOLITFTQELQDV 2137  
Qy 52 -----PEVKFNYYVDGVEVH----- 66  
Db 2138 VAKEKDTMATFECETSEFFVKVVKYKDGMEVHEGDKYRMHSDRKVHFLSILITIDTDAED 2197  
Qy 67 -----NAKT----- 70  
Db 2198 YSCVLVEDENVKTTAKLIVGAVVEFKELQDIEVPESYSGELECIYSPENIEGKWYHND 2257  
Qy 71 -----KPR----- 73  
Db 2258 VELKSNKYTTITRRGRQNLTKVDKREDOGEYSFVIDGKTTCKLKMKPRPIALQGLS 2317  
Qy 74 ----- 73  
Db 2318 DOKVCEGDIVQLEVKVLSLESVEGWMKDGQEVQPSDRVHIIVIDKOSHMLLIEDMTKEDAG 2377  
Qy 74 ----- 73  
Db 2378 NYSETIPALGLSTGRVSVYSDVITPLKDVNVIEGTAKVLECKVSPDVTSVKWLNDE 2437  
Qy 74 ----- 73  
Db 2438 QIKPDDRQVAIVKGTQKRLVINRTHASDEGYPKLIIVGRVETNCNLSVEKIKIIRGLRDLT 2497  
Qy 74 ----- 73  
Db 2498 CTETQNVVFEVLSHSGIDVLWNFKDKBKPKSKYKIEAHGKIYKLTVLNMKDDGKYT 2557  
Qy 74 ----- 73  
Db 2558 FYAGENMTSGKLTVAGGAISKPLTDQTVAESQEAFFEVEVANPDSKGEWLKDLPLTN 2617  
Qy 74 ----- 73  
Db 2618 NIRSBDGHRRLIIAATKLDIGETVYKVATSKTSAKLKVEAVKIKTKLNLVTFEQD 2677  
Qy 74 ----- 73  
Db 2678 AVFTVELTHPNVKGWQIKNGWLVLESNEKAIYSVKGTIYSLRKNAKNAIVDESIVGFRLGR 2737  
Qy 74 ----- 73  
Db 2738 LGASARLHVETVKTIKKPKDVTALENATVAFESVSHDTPVPKWFHKSVEIKPSDKHRLV 2797  
Qy 74 ----- 73

Db 2798 SERKVHKLMLQNISPDSAGEYTAVVQLECKAKFLVETLHITKTMKNIEVPETKTASPEC 2857  
Qy 74 ----- 73  
Db 2858 EVSHFNVPMSWLKNGVEIEMSEKFKIVVQGLHQLIIMNTSTEDSAEYTFVCGNDQVSAT 2917  
Qy 74 ----- 73  
Db 2918 LTVTPIMTSMKLDINAEKDTITFEVTVNYEGISYKWLKNGVEIKSTDKQMRTKKLTH 2977  
Qy 74 ----- 73  
Db 2978 SLNIRNVHFGDAADYTFVAGKATSTATLYVEARHIEFRKHIDIKVLEKKRAMFECEVSE 3037  
Qy 74 ----- 73  
Db 3038 PDITVQMMKDDOELQITDRIKIOKEYVHRLIIPSTRMSDAGKYTVVAGGNVSTAKLFE 3097  
Qy 74 ----- 73  
Db 3098 GRDYRIRSIKKEVQVIEKQRAVVEFEVNEDDVDAHMYKDGIEINFQVQERHKYVVERRIH 3157  
Qy 74 ----- 73  
Db 3158 RMFISETRQSDAGEYTFVAGNRSSVTLYVNAPEPPVQLQELQVPTVQSGKPARFCAMIS 3217  
Qy 74 -----EE----- 75  
Db 3218 GRPOPKISWYKEQLLSTGFKCKFLHDQGEYTLILLIEAFPEDAAVYTCEAKNDYGVATTS 3277  
Qy 76 ----- 75  
Db 3278 ASLSVEPEVYVSPQEMPVYPPAITPLQDVTVTSEGPQRCRVSGTDLKVSWSKDKK 3337  
Qy 76 -----QYNSTYRV----- 83  
Db 3338 IKPSRFRMTQFEDTYQLEIAEAYPEDEGTYTFVANNVAGVQSVSTANLSLEAPESILHER 3397  
Qy 84 ----- 83  
Db 3398 IEQIEIEMEMKFFSSFLSABEEGLHSAELQLSKINETLELLESVPYPTKFDSEKEGTGP 3457  
Qy 84 -----VSVL----- 87  
Db 3458 IFIKEVSNADISMGDVATLSVTVIGIPKPIQWPFNGVLLTPSADYKPFVDCDDHSLIL 3517  
Qy 88 ----- 87  
Db 3518 FTKLEDEGEYTCMASNDYGKTCISAYLKINSKGEGHKDTETESAVAKSLEKLGPCPPHF 3577  
Qy 88 -----TVLH----- 91  
Db 3578 LKELKPIRCAOGLPAIFEYTVVGEPAPTVTWFKENKQLCTSVYTIHHPNGSGTFIVND 3637  
Qy 92 ----- 91  
Db 3638 PQREDSGLYICKAENMLGESTCAEELLVLEDDTMDTPCKAKSTPEAPDPQTPKGP 3697  
Qy 92 ----- 91  
Db 3698 AVEALDSQEIATFVKDFTILKAALITEENQQLSYEHIKAKANELSSOLPLGAQLQSILEQ 3757  
Qy 92 ----- 91  
Db 3758 DKLTPESTREFLCINGSIHQFLKPESPNLQIIVQSQKTFKSKEGILMPEEPETOAVLSD 3817  
Qy 92 ----- 91  
Db 3818 TEKIPPSAMSTEQINSLTVEPLKTLLEAPEGNYPOSSIEPPMHSYLTSVAEVLSLKCKT 3877  
Qy 92 ----- 91

```
Db 3878 VSDTNRQVRTLOKQEAQSALILSQAEGHVESQSPDVMISQVNYEPLVPSEHSCTEG 3937
QY 92 ----- 91
Db 3938 GKILIESANPLENAGQDSAVRIEKGSLRPLALEEKQVLLKEEHSNDVMPDQIESK 3997
QY 92 -----ODWL- 95
Db 3998 REPVAIKKQVQGRDLLSKESLLSGIPEQRNLKIQICRALQAASQPGLFSEWLR 4057
QY 96 ----- 95
Db 4058 NIEKVEAVNITQEPHIMCYLVTSKASVTEEVTHIIBEDVPQMANLKMELRDALCAI 4117
QY 96 ----- 95
Db 4118 IYBEIDILTAEGPRIQOAGAKTSLQEMDSFGSQKVEPIPEPEVESKYLISTEVSFYFW 4177
QY 96 ----- 95
Db 4178 QSRVKYLDATPVTKGVASAVVDEKQDES LKPSKEESESSEGTBEVATVKIQEABGGL 4237
QY 96 ----- 95
Db 4238 IKEDGPMIHTPLVDTVSEGDIVHLTTSITNAKEVNMVNFENKLVPSDEKFKCLODQNTYT 4297
QY 96 -----NGK- 98
Db 4298 LVIDKVNTEDHQGEYVCEALNDSGKTATSAKLTJVVKRAAPVKKRTEPLEVALGHLAKFT 4357
QY 99 -----EYKCKVSN- 106
Db 4358 CEIQSAPNVRFWFKAGRIEYSDKCSIRSSKYISSLEILRTQVDCGEYTCASNEYGS 4417
QY 107 ----- 106
Db 4418 VSCATLTVTPGGEKKVRKLLPERPEKPEEVVLKSVLRKRPEEPEKPKLEKVKX 4477
QY 107 ----- 106
Db 4478 PAVPEPPPKPVEVEVPTVTKRERKIPEPTKVPEIKPAIPLPAPPEKPKPEAEVKTIP 4537
QY 107 -----KALPAPIE- 114
Db 4538 PVPEPTPTAAPTVPVVGKKAEBKAEAKPKGPIKGVPKKTPSPTEARRKLRPG 4597
QY 115 ----- 114
Db 4598 SGGEKPPDEAPFTYQLKAVPLKFEVKEIKDIIILTESEFVGSSEAFCELVSPSTAITWMKD 4657
QY 115 ----- 114
Db 4658 GSNIRESPKHRFTADGDKRKLHIIDVLSAGEYTCVLRGNKEKTS AKLVVEELPVRF 4717
QY 115 -KTS- 118
Db 4718 VKTLEBEVTVVGQPLYLSCELNKRDDVVVRKDKGKIIVBKPGRIVPGVIGLMRALTI INDA 4777
QY 119 ----- 118
Db 4778 DOTDAGTYTVVENANNLSCSCVKVVEVIRDLVLKPIRDQHVKPGKTAIFACDTAKDTP 4837
QY 119 ----- 118
Db 4838 NIKWFKGYDEIPAEPNDKTEILRDGNHLYLKIKNAMPEIDAEVAVIEGKRYPAKLTIGE 4897
QY 119 ----- 118
Db 4898 REVELLKPIEDVTIYEKESASFAEISEADIFQWMLKCELLRPSPTCEIKAEKGRFILT 4957
QY 119 ----- 118
Db 4958 LHKVKLDQAGEVLYQALNATITAILTVKEILDFAVPLKDVTPVRRQARFECVLTREAN 5017
```

---

```
QY 119 ----- 118
Db 5018 VIWSKGPDIKSSDKFDIIADGKKHILVINDSOFDDEGVYTAEEVGKTSARLFTVGTGIRL 5077
QY 119 ----- 118
Db 5078 KPMSPLEDOVTKEGETATFVCELSHEKMHVWFKNDAKLHSTRTVLISSEBKGTHKLEMKE 5137
QY 119 ----- 118
Db 5138 VTLDLDSIQAKQVKELSSAQKLVLEADPYFTVKLHDKTAVEKDEITLACEVSKDVPVKW 5197
QY 119 ----- 118
Db 5198 FKDGEIIVSPKYSIKADGLRLIKKADLKDKGVEYDCGTDKTKANVTVEARLIEVE 5257
QY 119 -----KAKQOP- 124
Db 5258 KPLYGVEVFGETAHFEIELSEPDVHGQWKLKGOPLTASPDCEIIEDGKKHILHNCOL 5317
QY 125 -----REPO- 128
Db 5318 GMTGEVSFOAANAKSAANLKVKEPLIFITPLSDVKVEKDEAKFECEVSEVPKTPRWLK 5377
QY 129 ----- 128
Db 5378 GTQEITGDDRFELIKDGTKHSMVKSAAFEDEAKYMFEAEDKHTSGKLIIEGIRLAKFLTP 5437
QY 129 ----- 128
Db 5438 LKDVTAKEKESAVFTVELSHDNIRVKWFKNDQRLHTRTSVSMODEGKTHSITFKDLSIDD 5497
QY 129 ----- 128
Db 5498 TSOIRVEAMGMSSEAKLTVLEGDPYFTGKLQDYTGVEKDEVILQCEISKADAPVKWFKDG 5557
QY 129 ----- 128
Db 5558 KEIKPSNAVIKTDGKKRMLILKALKSDIGQYTCDCGTDKTSGLKDIEDREIKLVRPLH 5617
QY 129 ----- 128
Db 5618 SVEVMEETARPETEISEDDIHANMKLGEALLQTPDCEIKEEGKTHSLVHLNCRLDQGTG 5677
QY 129 ----- 128
Db 5678 GYDFQAANYKSSAHLRVKPRVIGLLRPLKDVTVTAGETATFDCELSYEDIPVWYWLKGGK 5737
QY 129 ----- 128
Db 5738 LEPSDKVVPVPRSEGVHTLTRLRVKLEDAGEVOLTAQDFKTHANLFWKEPPEVFTKPLEDQ 5797
QY 129 ----- 128
Db 5798 TVEEGATAVLECEVSRENKAVKFKNGTEILSKYKYEIVADGVRKRLVIHDCPTPEDIKTY 5857
QY 129 ----- 128
Db 5858 TCDAKDFKTSCLNVVPPHVEFLRPLTDLOVREKEMARFECELSRENKAVKFKDGAETK 5917
QY 129 ----- 128
Db 5918 KGKKYDIISKGAVRILVINKCLDDDEAEYSCVARTARTSGMLTVLEEEAVFTKNLANIEV 5977
QY 129 ----- 128
Db 5978 SETDTIKLVCEVSKPAEVIWYKGDDEIETGTYEILTGRKRILVIQNAHLEDAGNYNC 6037
QY 129 ----- 128
Db 6038 RLPSSRTDGVKVKHELAAEFISKPONLEILEGEKAFFVCSISKESFPVQWKRDDKTLESG 6097
```

QY	129	-----	128
Db	6098	DKYDIADGKRVLVKDATLQDMGTVMVGAARAAHLTVIEKLRIVVPLKDRVKEQ	6157
QY	129	-----VTL-----	132
Db	6158	QEVVFNCEVNTGAKAKWFRNEEAIFDSSKYYIILOKDLVYTLRIRDAHLDDQANYNVSLT	6217
QY	133	-----	132
Db	6218	NHRGENVKSANLIVEBEDLRIVEPLKDIETWEKKSVTFWCKVNRNLNVTLKWTNGEVP	6277
QY	133	-----	132
Db	6278	FDRVSVYRVDKYKHLTIKDCGFPDEGEYIVTAGQDKSVAELIIIEAPTEFVEHLEQTV	6337
QY	133	-----	132
Db	6338	TEFDVAVFSCQLSREKANVWYRNGREIKEGKYKFEKDGSIHRLIIKDCRLDDECEYAC	6397
QY	133	-----	132
Db	6398	GVEDKSRARLFVEEIPVEIIRPPQDILEAPGADVFLAELNKDKVEQWLRNNWVVOG	6457
QY	133	-----	132
Db	6458	DKHQMSEGKIHLRIQICDIKPRDQGEYRFIAKDKEARAKLELAAAPKIKTADQDLVVDVG	6517
QY	133	-----	132
Db	6518	KPLTMVVPYDAYPKAAEWFENEPLSTKITDTTAEQTSFRILEAKKGDKGRYKIVLQNK	6577
QY	133	-----	132
Db	6578	HGKAEGFINLKVIVDPGVRNLEVTFETFDGEVSLAWEEPLTDGSKIIGYVYVRRDIKRK	6637
QY	133	-----	132
Db	6638	TWVLATDRAESCEFTVTGLQGGVEYLFVRSARNRVCTGEPVEIDNPVEARSKYDVPGP	6697
QY	133	-----	132
Db	6698	LNVTITDVRNFGVSLTWEPEYDGGAEITNYVIELRDKTSIRWDTAMTVRAEDLSATVTD	6757
QY	133	-----	132
Db	6758	VVEGQYSFRVRAQNRIGVGKPSAATPFVKVADPIERPSPVNLTSDDQTSVOLKWE	6817
QY	133	-----	132
Db	6818	PLKDGSPILGYIIERCEEKGNWIRCNMKLVPELTYKVTGLEKGNKYLYRVAENKAGV	6877
QY	133	-----	132
Db	6878	SDPSEILGLTADDAFVPTMDLSAFKDGLEVIPNPITILVPSTGYPRPTATWCFGDKV	6937
QY	133	-----	132
Db	6938	LETGDRVMKTL SAYAELVISPSERSDKGIYTLKLENRVKTIISGEIDVNVIA RPSAPKEL	6997
QY	133	-----	132
Db	6998	KFGDITKDSVHLTWEPPDDGGSPLTGYVVEKREVSRKWTWKYMDFTDLEFVTPDLVQG	7057
QY	133	-----	132
Db	7058	KEYLFKVCARNKCGEPAYVDEPNMSTPATVPDPENVKWRDRTANSIFLTWDPKND	7117
QY	133	-----	132
Db	7118	GGSRIGYIVERCPRGSKWVACGEPVAETKMEVTGLEGKWYAYRVKTLNRQGAKPSR	7177
QY	133	-----	132

Db	7178	PTEIQAVDTQEAPEIEFLDVKLLAGLTVKAGTKIELPATVIGKPEKITWTAKMILKQD	7237
QY	133	-----	132
Db	7238	KRITIENVPKSTVTIIVDSKRSDTGTYYII EAVNVCGRATAVVEVNVLDKPGPAAFDITD	7297
QY	133	-----	132
Db	7298	VTNESCLLTNPPRDDGSKITNYVVERRATDSEWHKLSSTVKDTNFKATKLIPNKEYI	7357
QY	133	-----PPSR---DELTKNOVSLT-----	147
Db	7358	FRVAENNYGAGEVQASPI TAKYQFDPGPGPTRLPSDITKDAVTLTWCEPDDDDGGSPI	7417
QY	148	-----	147
Db	7418	TGYVVERLDPDTDKWRCNKMPVKDTTYRVKGLTNKKYFRVLAENLAGPKPSKSTEP	7477
QY	148	-----	147
Db	7478	ILIKDPIIDPPWPGKPTVKDVGKTSVRLNWTKPEHDGGAKIESYVIELMKTGTDEWVRVA	7537
QY	148	-----	147
Db	7538	EGVPTTQHLLPGLMEGQEYSFRVRAVNKAGESEPSPDPVLCREKLYPPSPPRWLEVIN	7597
QY	148	-----	147
Db	7598	ITKNTADLKWTVPEKGGSPITNYIVEKRDVRRKGWQTVDTTVKDTKCTVPLTEGSLYV	7657
QY	148	-----	147
Db	7658	FRVAENNAIGOSDTEIETEDSVLAKDFTTTPGPPYALAVDVTKRHVDLKWEPPKNDGGRP	7717
QY	148	-----	147
Db	7718	IQRVYIEKKERLGRTRWKAGKTAGDCNFRVTDVIEGTEVOFQVRAENEAGVHPSEPTE	7777
QY	148	-----	147
Db	7778	ILSIEDPTSPSPPLDLHLVTDAGRKHIAIAWKPKPEKNGSPIIGYHVEMCPVGTEKWMRV	7837
QY	148	-----	147
Db	7838	NSRPIKDLKFKVEGVVDPDKEYLVRRAVNAIGVSEPSEISENVVAKDPCKPTIDLETH	7897
QY	148	-----	147
Db	7898	DIIVIEGEKLSIPVPFRAVPVPTVSHKDGKEVKASDRLTMKNDHISAHLEVPKSVRADA	7957
QY	148	-----	147
Db	7958	GIYTTITLENKLSATASINVKVIGLPGPCKDIKASDITKSSCKLTWEPPEPDGPTILHY	8017
QY	148	-----	147
Db	8018	VLERREAGRRTYIPVMSGENKLSWTVDLIPNGEYFFRVKAVNKVGGEYIELKNPVIAQ	8077
QY	148	-----	147
Db	8078	DPKQPPDPVDEVHNPTAEAMTITWKPLYDGGSKINGYIIIEKIAKEERWKRCKNEHLV	8137
QY	148	-----	147
Db	8138	PILTYTAKLEEGKEYQFRVRAENAAIGSEPSRATPTTAYDPIADAPKVIILRTSLEVRKG	8197
QY	148	-----	147
Db	8198	DEIADASISGSPYPTITWIKDENVIVPEEIKKRAAPLVRRRRKGVEEPEFVLPLORL	8257
QY	148	-----	147

Db 8258 SIDNSKGSQSLRSLRPHDGLYMIKVENDHGIAKAPCTVSVLDTGPPINFVFEDIR 8317  
QY 148 ----- 147  
Db 8318 KTSVLCKWBPPLDDGSEIINTLEKKDKTKPDSEHIVVTSTLRHCKYSVTKLIEGKEYL 8377  
QY 148 ----- 147  
Db 8378 FRVRAENRFGPGPCVSKPLVAKDPGPDADPKPIVEDVTSNSMLVKWNEPKDNGSPLL 8437  
QY 148 ----- 147  
Db 8438 GYWLKREYNSTHWSRVNKSLLNALKANYDGLLEGLTYFVRVCAENAAAGPKFSPDPK 8497  
QY 148 ----- 147  
Db 8498 TAHDPISPFGPPIPRVTDTSSTIELEWEPFAENGGEIVGVFDVKQLVGTNKNRSCTEK 8557  
QY 148 ----- 147  
Db 8558 MIKVRQYTVKEIREGADYKLRVSANVNAAGEPPGETQPVTVAEPPQEPAPVELDVSKGI 8617  
QY 148 ----- 147  
Db 8618 QIMAGKTLRIPAVVTGRPVPTKVTWTKKEGELDKDRVVIDNVGTSKSELIITKDALKRDHGRY 8677  
QY 148 ----- 147  
Db 8678 VITATNSCGSKFAARVEFDVPGPVLDLKPVVTNRKMKLLNWSDEDDGSGSEITGFIIE 8737  
QY 148 ----- 147  
Db 8738 RKDAKHTWRQPIETERSKCDITGLLEGQYKFRVIAKNKFCGPPVEIGPILAVDPLGP 8797  
QY 148 ----- 147  
Db 8798 PTSPEBLTYTERQSTITLDWKEPRNSGSPIQGYIIERRRHKPDFERVNKRCLCPTTSF 8857  
QY 148 ----- 147  
Db 8858 LVENLDHQMYEYFRVKAUNEIGESEPSLPLNVVQDDEVPTIKLRLSVRGDTIKVKAGE 8917  
QY 148 ----- 147  
Db 8918 PVHIPADVTGLPMPKIEWSKNETVIEKPTDALQITKEEVSSEAKTELSIPKAVREDKGT 8977  
QY 148 ----- 147  
Db 8978 YVTASNRLGVSFRNVHVEYDRPSPRNLAVIDIKAESCYLTWDAPLNGGSEITHYVI 9037  
QY 148 ----- 147  
Db 9038 DKRDASRKAWEVNTAVERKYGWIWKLIPNGQYEFVRVAVNKYISDECKSKVVIQD 9097  
QY 148 ----- 147  
Db 9098 PYRLPGPGCKPKVLARTKGSMLVSWTPTPLDNGSPITGYWLEKREBGSFYWRSRAPIT 9157  
QY 148 ----- 147  
Db 9158 KVGLKGVFNVPRLLEGVKYQFRAMAINAAGIGPPSPDPVAGDPIPPPGPPSCPEVK 9217  
QY 148 ----- 147  
Db 9218 DKTKSSISLGWPPAKDGGSPKIGYIVEMQEEGTTDKWRVNEPDKLIITTCBVCVWPNLKE 9277  
QY 148 ----- 150  
Db 9278 RKYFRVKAUNEAGESEPSDTTGEIPATDIOEPEVFDIGAQCCLVCKAGSQIRIPAVI 9337  
QY 151 ----- 150  
Db 9338 KGRPTPKSWEFDGKAKAMKDVHIDPAQLETAENSVIIPECKRSHTGKYSITAK 9397

QY 151 -----KGF----- 153  
Db 9398 NKAGOKTANCYKVMVDPGPPKDLKVS DITRGSCRLSWKMPDDGGDRIKGVIEKRTID 9457  
QY 154 -----YPSD- 157  
Db 9458 GKAWTKVNPDCGTTTFVVPDLLEQOYFPRVRAENRFGIGPPVETIQTARTARDPIYPPDP 9517  
QY 158 ----- 157  
Db 9518 PIKLIKIGLITKNTVHLSWKPKNDGGSPVTHYIVECLAWDPTGTKEAWROCNRKRDVEEL 9577  
QY 158 ----- 157  
Db 9578 QFTVEDLVEGGYEFVRVKAUNAAGVSKPSATVGPCDCQCPMDPPSIDLKEFMEVEEGTNV 9637  
QY 158 ----- 157  
Db 9638 NIVAKIKGVPPPTLTWFKAPKKPDNKEPVLVDTHVNKLVDCTCTLVIPQSRSDTGLY 9697  
QY 158 ----- 157  
Db 9698 TITAVNNLGTASKEMRLNVLGRPPVGP IKFESYSAQMTLSWPPKDDGSGKITNYVI 9757  
QY 158 ----- 157  
Db 9758 EKREANRKTWVHSSEPKECTYTI PKLLEGHEYVFRIMAONKYGIGELDSEPETARNLF 9817  
QY 158 -----IAVEWE----- 163  
Db 9818 SVPGAPDKPTVSSVTRNSMTVNWEEPEYDGGSPVTGYWLEMKDTSKRWKRVNRDPIKAM 9877  
QY 164 ----- 163  
Db 9878 TLGVSYKVTGLIEGSDYQFRVYAINAAGVGPASLPSDPATARDPIAPGPPPKVTDWTK 9937  
QY 164 ----- 163  
Db 9938 SSADLENSPPLKDGSKVTGYIVEYKEBGEKEWKGKDEVRGTKLVTGLKEGAFYKFR 9997  
QY 164 ----- 163  
Db 9998 VSAVNIAGIGEPEVTDVIEMKDRLVSPDLQDASVRDRIVVHAGGVIRIIAYVSGKPPP 10057  
QY 164 ----- 163  
Db 10058 TVTWNMERTLPQEATIIETTAISSMWIKNCORSHOGVYSLAKNEAGERKKTIIVDVL 10117  
QY 164 -----SNGOPENNY----- 172  
Db 10118 VPCPVGTPTFLAHLNLTNESCCLTWFSPEDDGSPITNYVIEKRESDRRAWTPVYTYTTRQN 10177  
QY 173 ----- 172  
Db 10178 ATVGLIQGKAVFPRIAAENSIGMPFVETSEALVIREPITVPERPEDLEVKEVTKNTWT 10237  
QY 173 ----- 172  
Db 10238 LTWNPPKYDGGSEIINYVLESRLIGTEKFHKVTNDNLLSRKTYVKLKEGDTYEYRSVAV 10297  
QY 173 ----- 172  
Db 10298 NIVGOGKPSFCTKPTCKDELAAPTTLHLOFRDKLTIRVGEAFALTGRYSGRKPVKVSWFK 10357  
QY 173 -----KTP----- 176  
Db 10358 DEADVLEDDRTHIKTTPATLALEKIKAKRSDSGKYCVVVENSTGSRKGCQVNVVDHGP 10417  
QY 177 ----- 176  
Db 10418 PVGPVSFDEVTKDYMVISWKPPLDDGSGKITNYIIIEKKEVGKDVMPVTSASAKTTCKVS 10477

QY 177 ----- 176  
Db 10478 KLEGGDYIFRIHAENLYGISDPLVSDSMKAKDRFRVPADQPIVTEVTKDSALVTWNK 10537  
QY 177 ----- 176  
Db 10538 PHDGGKPTITNYILEKRETKSRWARVTKDIHPYTKFRVPDLLEGGQYEFVSAENEIGI 10597  
QY 177 ----- 176  
Db 10598 GDSPPSKVPFAKDPIAKSPVPVNPDAIDTTCNSVDLTWQPPRHDDGSKILGYIYEQV 10657  
QY 177 ----- 176  
Db 10658 GDEWRRAHTPESCPETKYKVTGLRDGQYKFRVLAVNAAGESDPAHVPEPVLVKDRLE 10717  
QY 177 ----- 176  
Db 10718 PPELILDANMAREQHIVGDTLRLSAIIKGVPPFKVTWKEDRDAPTARIDVTPVGSKL 10777  
QY 177 ----- 176  
Db 10778 EIRNAAHEDGGIYSLIVENPAGSKTVSVKVLVDKPGPPRDLEVEIRKDSCHLTWKEPL 10837  
QY 177 ----- 176  
Db 10838 DGGSVITNVVERDVASQAQWSPLSATSKKSHFAKHLNEGNOYLFRVAENQYGRGPF 10897  
QY 177 ----- 176  
Db 10898 VETPKIKALDPLHPGPKDLHHVDVKTEVSLVMNKPDRDGGSPITGYLVEYQEBGTQ 10957  
QY 177 ----- 176  
Db 10958 DWIKFKVTNLECVVTGLQGGKYRFRVKAENIVGLGLPTTPIECQEKLVPPSVELDV 11017  
QY 177 ----- 176  
Db 11018 KLIEGLVWKAGTVRRPAIIIRGVPTAKWTTDGESEIKTDEHYTVETDNFSSVLTIKNCL 11077  
QY 177 ----- 176  
Db 11078 RRDTEGYQITVSAAGSKTVAVHLTVLDVPGPPTGINILDVTPHEMTISWQPPKDDGGS 11137  
QY 177 ----- 176  
Db 11138 PVINIVEKQDTRKDTGWVSSGSSKTKLIPHQKCEYVFRVRAENKICVGPPLDSTP 11197  
QY 177 ----- 176  
Db 11198 TVAKHKFSPSPGKPVVTDITENAATVSWTLPKSDGSGPITGYIMERREVTGKWRVVK 11257  
QY 177 ----- 176  
Db 11258 TPIADLKFRVTGLYEGNTYEFVFAENLAGLSKPSSPDIKACRPKPPGPPINPKLD 11317  
QY 177 ----- 176  
Db 11318 KSBETADLVTKPLSDGSPILGVVVEQKPGTAQWNRINKDELIROCAFRVPGLTIEGNE 11377  
QY 177 ----- 176  
Db 11378 YFRIRAAINIVGEGPRELAESVIAKDILHPPEVELDVTCDVITVRVGOTIRILARVKG 11437  
QY 177 ----- 176  
Db 11438 RPEPDITWTKGKVLVREKRVLDLIQDLPRVELOIKEAVRADHGKYIISAKNSSCHAQGSA 11497  
QY 177 ----- 176  
Db 11498 IVNVLDPRPGCONLKVNTVTKENCTISWENPLDGGSEITNFIVEYRKNQKGSIVASD 11557  
QY 177 ----- 176

Db 11558 VTKRIKANLANNEYFRVCAENKVGVPETIETKTPILAINPIDRPEPENLHIADKKG 11617  
QY 177 ----- 176  
Db 11618 TFVYLKWRRPDYDGGSPNLSYHVERRLKGSDDWERVHKSGIKETHYMVDRCVENQIYEF 11677  
QY 177 ----- 180  
Db 11678 VQTKNEGSDWVKTBEVVVKEDLQKPVLDLKLGSVLTVKAGDTIRLEAGVRGKPPPEVA 11737  
QY 181 ----- 193  
Db 11738 WTKDKDATDLTRSPRVKIDTRADSSKFSKAKRSDGGKYVVTATNTAGSFVAYATVNV 11797  
QY 194 ----- 193  
Db 11798 DKPGPVNRLKIVDVSSDRCTVCWDPPEDDGGCEIQNYILEKCTKRMVWSTYSATVLTGP 11857  
QY 194 ----- 195  
Db 11858 TTVTRLIEGNEYIFRVAENKIGTPTESKPVIAKTKYDKPGPDPEVTKVSKBEMTV 11917  
QY 196 ----- 195  
Db 11918 VWNPEYDGGKSTGYFLEKKEKHSTRWPVNKSAIPERRMKVQNLLPDHEYQFRVKAEN 11977  
QY 196 ----- 195  
Db 11978 EIGEPESLPSRPVAKDPIEPGPPTNFRVDTTKHSITLGMGKPYDGGAPIGYVVE 12037  
QY 196 ----- 195  
Db 12038 MRPKIADASDEGWKRCNAAAQLVRKEFTVTSIDENQYEFVCAQNOVGIGRPAELKEA 12097  
QY 196 ----- 195  
Db 12098 IKPKEILEPPEIDLDASMRKLIVIRAGCPIRLFAIVRGRPAKVTWRKVGIDNVVRKGQV 12157  
QY 196 ----- 195  
Db 12158 DLVDTMAFLVIPNSTRDSDSKYSLTLVNPAGEKAVFVNVVRVLTDPGVSCLKVSDVTKTS 12217  
QY 196 ----- 195  
Db 12218 CHVSWAPPENDGSGQYTHYIVKREADRKTWSTVTPVEKTSFHVTLNLPVPGNEYFRVTA 12277  
QY 196 ----- 195  
Db 12278 VNEYGCVPTDVPKPVILASDPLSEPDPPRKLATEMTKNSATLAWLPLRDLGGAKIDGYI 12337  
QY 196 ----- 195  
Db 12338 ISYREEQPADRWTEYSVVKDLSLVVTGLKEGKKYKFRVAARNAVGSVLPREAGVYEAK 12397  
QY 196 ----- 202  
Db 12398 EQLLPKILMPEQITIKAGKKLRIEAHYGKPHPTCKWKKGEDEVVTSSSHVAHVHADSSS 12457  
QY 203 ----- 202  
Db 12458 ILIIKDVTBKDSGYISLTAENSSGTDQKIKVVVMDAPGPPQPFDISIDADACSLSWH 12517  
QY 203 ----- 202  
Db 12518 IPLEDGGSNTINIVEKCDVSRGDMWTALASVTKTSCRVKLIPGOEYIFRVAENRFGI 12577  
QY 203 ----- 202  
Db 12578 SEPLTSPKVAQFPFGVPSFKNARVTKVNKDCIFVANDRPDSGSGPIIGYLIERKERN 12637  
QY 203 ----- 202

Db 12638 SLLWVKANDTLVRSTEYPCAGLVEGLEYSFRITYALNKGAGSPSPKTEYVYTARMVPDPG 12697  
QY 203 ----- 202  
Db 12698 KPEVIDVTKSTVSLIWARPKHDGSKIIGYFVEACKLPDGKWKVRCNTAPHQIPOEEYAT 12757  
QY 203 ----- 202  
Db 12758 GLEEKAOYQFRAIARTAVNISPPSPDPVTILAENVPPRIDLSVAMKSLLTVKAGTNCV 12817  
QY 203 ----- 202  
Db 12818 LDATVFGKPMPTVSMKDKDTLLKPAEGIKMAMQRLCTLELFSVNRKDSGDYTTAENS 12877  
QY 203 ----- 202  
Db 12878 GSKSATIKLVLDKPGPPASVKINKMYSDRAMLSWEPPLDGGSEITNIVDKRETSRPN 12937  
QY 203 -----VFSCV----- 208  
: ||||  
Db 12938 WAQVSATVPITSCSEKLEHGHGHEYOFRICAENKYGVDVFTPEPAIAKNPYDPPGRCDDP 12997  
QY 209 ----- 208  
Db 12998 VISNITKDHMTVSWKPPADDGGSPITGYLLEKRETOAVNWKVNRKPIERTLKATGLQE 13057  
QY 209 ----- 208  
Db 13058 GTEYEFVTAINKAGPKPSDASKAAYARDPOYPAPPAPFKVYDTTRSSVSLSWGKPAY 13117  
QY 209 ----- 208  
Db 13118 DGGSPILGYLVEBKRADSDNWVRCLPQNLQKTFEVTGLMEDTOYQFRVAVNKGISD 13177  
QY 209 ----- 208  
Db 13178 PSDVPDKHYPKDILIPPEGEHDADLRKTLILRAGVTMRLYVPVKGRPPPKITWSKPNVL 13237  
QY 209 ----- 208  
Db 13238 RDRIGLDIKSTDFDLRCENVNKYDAGKYILTLNENSGKKEYTIVVKVLDTPGPPINVT 13297  
QY 209 ----- 208  
Db 13298 VREISKDSAYVTWEPPIIDGGSPILNYYVQKRDAERKSWSTVTTECSKTSFRVFNLEBG 13357  
QY 209 ----- 208  
Db 13358 SYFFRVAENEYIGDGPGETRDAVKASQTPGPVVDLKVRSVSKSSCSIGWKKPHSDGSR 13417  
QY 209 ----- 208  
Db 13418 IIGYVVDLFTENKQVRNWKSLSQYSAKDLTEGKEYTFRVSAENEGEGTPEITVVAR 13477  
QY 209 ----- 208  
Db 13478 DDWAPDLDLKGLPDLCLAKENSFRLKIPIKGPAPSVSWKKGEDPLATDTRSVESS 13537  
QY 209 ----- 208  
Db 13538 AVNTTLIVDCQSDAGKYTTILKNVAGTKEGTISIKVYKGPITGPIKFDEVTAEAMT 13597  
QY 209 -----MHEALH----- 214  
: |||  
Db 13598 LKWAPKDDGGSEITNILEKRDVNNKWTVCASAVQKTTFRVTLHEGMEYTFRVSAEN 13657  
QY 215 ----- 214  
Db 13658 KYGVGGLKSEPIVARHPFDVPDAPPPNIVDRHDSVSLWTDPKKTGGSPITGYHLEF 13717  
QY 215 ----- 214  
Db 13718 KERNSLLWKRANKTPIRMDFKVTGLTEGLEYEFRVMAINLAGVKPSLPSEPVALDPI 13777

QY 215 ----- 214  
Db 13778 DPPGKPEVINITRNSVTLLIWTPEPKYDGGHKLGTGYIVEKRDLPKSKWMKANHVNPECAFT 13837  
QY 215 ----- 214  
Db 13838 VTDLVEGGKYEPRIRAKNTAGAISAPSESTETIICKDEVEAPTIVLDPTIKDGLTIKAGD 13897  
QY 215 ----- 214  
Db 13898 TIVLNAISILGKPLKPSKSKAGKDIRPSDITQITSTPTSSMLTIKYATRKDAGEYTTITA 13957  
QY 215 ----- 214  
Db 13958 TNPFGTKVEHVKVTVLDVPGPPGPVEISNVSAEKATLTWTPPLEDGGGSPKSIYILLEKRET 14017  
QY 215 -----NH----- 217  
: |||  
Db 14018 SRLLTWVVEDIQSCRHVATKLIQOGNEYIFRVSANVHYKGEPVQSEPVKMWDRFGPPG 14077  
QY 218 ----- 217  
Db 14078 PEKPEVSNVTNTATVSWKRPVDDGGSEITGYHVERREKSLRWVRAIKTPVSDLRCKVT 14137  
QY 218 -----TQKSLSLS----- 225  
: |||  
Db 14138 GLOBGSTYEFVRVSAENRAGIGPPSEASDVLMDAAYPPGPPSPNPHVTDTKKSASLAWG 14197  
QY 226 ----- 225  
Db 14198 KPHVDGGLTGYVVEHQKVGDEAWIKDTGTALRITQFVVPDLQTKKYNFRISAINDA 14257  
QY 226 ----- 225  
Db 14258 VGGEPAVIPDVEIVEREMAPDFELDAELRRTLVRAGLSIRIFVPIKGRPAPEVTWTDN 14317  
QY 226 ----- 225  
Db 14318 INLKNRANIENSTFTLLIIPECNRYDTGKFWMTIENPAGKSGFVNVRVLDTPGPVNL 14377  
QY 226 ----- 225  
Db 14378 RPTDITKDSVTLHWDPLIDGGRITNIVEKREATRSYSTATTCKHCTKYVTGLSEG 14437  
QY 226 ----- 225  
Db 14438 CEYFFRVAENEYIGIGEPTETTEPVKASEAPSPDLSNIMDITKSTVSLAWPKPKHDGS 14497  
QY 226 ----- 225  
Db 14498 KITGVVTEAQRKSGDQWTHITTVKGLECVVRNLTEGEYTFQVMVNSAGRSAPRESRPV 14557  
QY 226 ----- 225  
Db 14558 IVKEOTMLPELDLRGIYQKLVIAKAGDNKIVEIPVLRPKPTVTWKKGDQILKQTVRNF 14617  
QY 226 ----- 225  
Db 14618 ETTATSTILNTNECVSRSDSGPYPLTARNIVEGVGDVITIQVHDIPGPTGPIKFDEVSSD 14677  
QY 226 ----- 225  
Db 14678 FVTFSWDPENDGGVPISNYVVMRQDSTTWELATTVIRTTYKATRLTGLLEYQFRVK 14737  
QY 226 ----- 225  
Db 14738 AQNRVYGVGPGITSAMIWVANYPFKVPDPPTQVAVTKDSMTISWHEPLSDGSPILGYH 14797  
QY 226 ----- 225  
Db 14798 VERKERNGLMOTVSKALVPGNIFKSSGLTDGIAIEFRVIAENMAGKSKPSKSEPMLAL 14857

QY	226	-----	225
Db	14858	DPIDPGKPVPLNTRHTVTLKWKAPETGFGKITSYIVEKRDLPNGRWLKNFSLNLEN	14917
QY	226	-----	225
Db	14918	EFTVSGLTDAAYEFVRVIAKNAAGAISSPSEPSDAITCRDDVEAPKIKVDVKFKDTVLK	14977
QY	226	-----	225
Db	14978	AGEAFRLADYSCRPPPTMEWSKDGKELEGTAKLEIKIADFSTNLVKNKSTRRDSGAYTL	15037
QY	226	-----	225
Db	15038	TATNPGGFAKHFNKVLDRPGPEGPLAVTEVTSEKCVLSWFPPLDDGGAKIDHVIQK	15097
QY	226	-----	225
Db	15098	RETSRLAWTNVASEVQTKLVTKLLKGNEYIFRVMVANKYGVGEPLSEPVLAVPNYPG	15157
QY	226	-----	225
Db	15158	PDPKPNKPVTTITKDSMVVCHGHPDSDGSGSEIINYIVERRDKAGQWIKCNKKTLDLRY	15217
QY	226	-----	225
Db	15218	KVSGLTEGHEYEPRIMAENAGISAPSPTPFYKACDTPFKPGPGNPRVLDTSRSSISI	15277
QY	226	-----	225
Db	15278	AWNKPIYDGGSEITGYMVEITAEPEDEWQIVTPPAGLKATSYITITGLTENQEKIRIYAM	15337
QY	226	-----	225
Db	15338	NSEGLGEPALVPCTPKAEDBMLPEIELDADLRKVVTIRACCTLRLFVPIKGRDPPEVKW	15397
QY	226	-----	225
Db	15398	ARDGESLDKASTESASYTLIIIVGNVRFDGKYILTVENSSGSKSAFVNRVLDTPGP	15457
QY	226	-----	225
Db	15458	PQDLKVKVTKTSVTLTWBPPLDGGSKIKNYIVEKRESTRKAYSTVATNCHTKSWKVDQ	15517
QY	226	-----PGK-----	228
Db	15518	LQEGCSYIFRVLAENEYGIGLPAETAESVKASERPLPPGKITLMDVTRNSVSLSEKPEH	15577
QY	229	-----	228
Db	15578	DGGSRIILGYIVEMQTKGSKWATCATVKVTEATITGLIOGEYSFRVSAQNEKGISDPRQ	15637
QY	229	-----	228
Db	15638	LSVPVIAKDLVIPAPKLLFNFTVLAGEDLVDPFPFGRPTPAVTHHKDNVPLKQTRV	15697
QY	229	-----	228
Db	15698	NAESTENNSLLTIKDACREDVGHVVVKLTNSAGEAIELTNVIVLDKPGPTGPKMDEVIT	15757
QY	229	-----	228
Db	15758	ADSIITLSWGPYPKYDGGSSINNYIVEKRDSTTTWQIVSATVARTTIKACRLKTGCEYQFR	15817
QY	229	-----	228
Db	15818	IAAENRYGKSTYLNSEPTVAQYKVPKPGPTPVVTLSSRDSMEVQWNEPISDGSRVIG	15877
QY	229	-----	228
Db	15878	YHLERKERNILWVKLNKTPIPQTKFTGLEGEVEYEFVRVSAENIVGKPKSKVSECYV	15937
QY	229	-----	228

Db	15938	ARPDCCDPGRPEAIIVTRNSVTILQWKKPTYDGGSKITGYIVEKKELPEGRMMKASFTNII	15997
QY	229	-----	228
Db	15998	DTHFEVTGLVEDHRYEFVRVIAARNAAGVFSPESESTGAITARDEVDPPRISMDPKYKDTIV	16057
QY	229	-----	228
Db	16058	VHAGESFKVDADIYKPIPTIOWIKGOELSNRTARLEIKSTDFTATSLSVKDAVRVDSGNY	16117
QY	229	-----	228
Db	16118	ILKAKNAVAGERSVTNVKVLDRPPEGPVVISGVTAEKCTLAWKPLQDGGDIINYIV	16177
QY	229	-----	228
Db	16178	ERRETSRLVMTVVVDANVQTLSCVKTKLLEGNEYTFRIMAVNKYGVGEPLSEPVVAKNPF	16237
QY	229	-----	228
Db	16238	VVPDAPKAPVTTVTXDSMIVVWERPASDGSSEILGYVLEKRDKEGIRWTRCHKRLIGEL	16297
QY	229	-----	228
Db	16298	RLRVTGLIENHDYEFVRVSAENAGLSEPPSPAYQACDPIYKPGPNPKVIDITRSSV	16357
QY	229	-----DWLKAFTD-----	236
Db	16358	FLWSKPIYDGGCEIOGYIVEKCDVNVGEMTCTPTGTINKNTIEVEKLEKHEYNFRIC	16417
QY	237	-----	236
Db	16418	AINKAGVGEHADVPPIIVEEKLPAIDLDLRLKIINIRAGSLSLFLVPIKGRPTPEV	16477
QY	237	-----	236
Db	16478	KMGKVDGEIRDAAIIDVTSFTSLVLDNVNRYDSGKYTLTLENSSGTKSAFVTVRLDTP	16537
QY	237	-----	236
Db	16538	SPVNLKVTEITKDSVTSITWEPPLDGGSKIKNYIVEKREATRKSAAVVTNCHKNSWKI	16597
QY	237	-----KVAE-----	240
Db	16598	DOLQEGCSYIFRYTAENEYIGLPAQTAADPIKVAEVPQPPGKITVDDVTNSYSLSWTKP	16657
QY	241	-----	240
Db	16658	EHDGSKIIQYIVEMOAKHSEKSEKARVKSLOAVITNLTQGBEYLFRRVAVANEKGRSDP	16717
QY	241	-----	240
Db	16718	RSIAVPIVAKDLVIEPDKPAFSSYSVQVQDLMKEVPIISGRPKPTITWTKDGLPLKQTT	16777
QY	241	-----	240
Db	16778	RINVTDSLDTLTLISIKETHKDDGGQYITVANVVGOKTASIEIVTLDKPDPKPGVKFDD	16837
QY	241	-----	240
Db	16838	VSAESITLSWNPPLYTGCCQITNVIQKRDTTWTVDVVSATVARTTLKVTCLKTGTEYQ	16897
QY	241	-----	240
Db	16898	FRIFAENRYGQSFALSDPIVAQYKYKEPGPGPTPFATAISKDSMWIQWHEPVNNGSPV	16957
QY	241	-----	240
Db	16958	IGYHLERKERNILWTKVKNKTIHDTQFKAQNLBEGIEYEFVRVSAENIVGVGKASNSEC	17017
QY	241	-----	240

Db 17018 YVARPCDPPGTPEPIWKRNEITLQWTKPVYDGGMITGYIVEKRDLPDGRMKASFTN 17077  
QY 241 ----- 240  
Db 17078 VIETQFTVSGLTEDQRYEPRVIAKNAAGSAIKSPSTGPIITAKDEVLPRIKMDPKFRDT 17137  
QY 241 ----- 240  
Db 17138 IVVNAGETFRLEADVHGRPLPTIENLGRDKEIEESARCEIKNTDFKALLIVKDAIRDGG 17197  
QY 241 ----- 240  
Db 17198 QVILRASNVAGSKSPFVNVKVLDRPPEGPVQVGTGVTSEKSLTWSPLQDGGSDISHY 17257  
QY 241 ----- 240  
Db 17258 VVEKRETSRLAWTVASEVVTNSLKVTKLLEGNVYFRIMAVNKYGVGEPLSAPVLKMN 17317  
QY 241 ----- 240  
Db 17318 PFVLPGPKSLEVTNIANDSMIVCHNRPSDGGSEIIGYIVEKRDORSIGIRWIKCNKRIT 17377  
QY 241 ----- 240  
Db 17378 DLRLRVTLGTEDHEYEFVSAENAAAGVGEPSPATVYKACDPVKPGPPTNAHIVDTTKN 17437  
QY 241 ----- 240  
Db 17438 SITLAWKPIYDGGSEIIGVVEICKADEEWOIVTQTLRVTRFEISKLTEHOEYKIR 17497  
QY 241 ----- 240  
Db 17498 VVALNKVGLGEATSVPGTVKPEDKLEAPELDDSELKRGIVVRAGSARIHIPFKGRPMP 17557  
QY 241 ----- 240  
Db 17558 EITWSREEGEFTDKVQIEKGVNYTOLSIDNCDRNDAGKYILKLENSGSKSAFVTVKVL 17617  
QY 241 ----- 240  
Db 17618 TPGPPONLAVKVRKDSAFVWEPPIIDGGAKVKNYVIDKRESTRKAYANVSSCKSTSF 17677  
QY 241 ----- 240  
Db 17678 KVENLATEGAIYFRVMAENEFVGVVETVDAVKAAPPPGKVTLDVDSQTSASLWME 17737  
QY 241 ----- 240  
Db 17738 KPEHGGSRVLGYVVMQPKTEKWSIIVAESKVCNAVVTGLSSGOEYQFRVKAYNEKGS 17797  
QY 241 ----- 240  
Db 17798 DPRVLGVPVIAKDLTIQPSLKLFPNTYSIQAGEDLKIPIVIGRPRPNISWVKDEPLKQ 17857  
QY 241 ----- 240  
Db 17858 TTRVNVEETATSTVLHIKEGKNDGKVTATNSAGTATENLSVIVLEKPGPPVGPVRF 17917  
QY 241 ----- 240  
Db 17918 DEVSADFVVISWEPPAYTGGCOISNYIVEKRDTTTTTHWMSATVARTTIKITLKTGTE 17977  
QY 241 ----- 240  
Db 17978 YQFRIFAENRYKSAPLDSKAVIVQYPFKEPDPGTPPTVTSIKDQMLVQWHEPVNDGGT 18037  
QY 241 ----- 240  
Db 18038 KIIGYHLEQKEKNSILWVKNLKTPIQDTKFKTTGLDEGLEVEFKVSAENIVGIGKPSKVS 18097  
QY 241 ----- 240  
Db 18098 ECFVARDPCDPPGREPAIVITRNNVTWKWKKPAYDGGSKITGYIVEKRDLPDGRMKASF 18157

QY 241 ----- 240  
Db 18158 TNVLETEFTVSGLVEDQRYEPRVIAARNAAGNFSEPSDSSGAITARDEIDAPNASLDPYK 18217  
QY 241 ----- 240  
Db 18218 DVIVVHAGETFVLEADIRGKPIPDVVMVSKDGKLELETAARMEIKSTIQKTTLLVWKCIRT 18277  
QY 241 ----- 240  
Db 18278 DGGQVILKLSNVGTSIPITVKVLDLRPGSPGPKLVGTVAEKCYLAWNPLODGGANI 18337  
QY 241 ----- 240  
Db 18338 SHYIIEKRETSRLSWTQVSTEVOALNWKVTKLLPGNEYIFRVMVKNKYGIGEPLESGPVT 18397  
QY 241 ----- 240  
Db 18398 ACNPYKPPGPPSTPEVSAITKDSMVVVTWARVPDDGGTEIEGYILLEKRDKEGVRWTKCNK 18457  
QY 241 ----- 240  
Db 18458 TLTDLRLRVTLGTGHSYEFVRAAENAAAGVGEPSVYFACDALYPPGPPSNPKVTD 18517  
QY 241 ----- 244  
Db 18518 SRSSVSLAWSKPIYDGGAPVKGYVVEKAAADEWTTCTPPTGLQGQFTVTKLKENY 18577  
QY 245 ----- 244  
Db 18578 NFRICAINSEGVGEPATLPGSVVAQERIEPEIELDADLRKVVLRASATLRLFTIKGR 18637  
QY 245 ----- 244  
Db 18638 PEPEVKWEKAGILTDRAQIEVTSFTMLVIDNVTFRDGRYNTLENNSGSKTAFNVNR 18697  
QY 245 ----- 244  
Db 18698 VLDSPAPVNLTIREVKKDSVTLSEPPPLIDGGAKITNYIVEKRETRKAYATITNNCTK 18757  
QY 245 ----- 244  
Db 18758 TTFRIENLOEGCSYFRVLASNEYGIGLPAETTEBPVKVSEPPLPBGRVTLVDVTRNTATI 18817  
QY 245 ----- 244  
Db 18818 KWKEPSDGGSKITGYVVMQTKGSEKWTCTQVKTLEATISGLTAGEEYVFRVAANVEK 18877  
QY 245 ----- 244  
Db 18878 GRSDPRQLGVPVIAARDIEIKPSVELPHTFNVKAREQLKIDVPFKGRPOATVNNRKDQOT 18937  
QY 245 ----- 244  
Db 18938 LKETTRVNVSSKTVTSLSIKEASKEDVGYELCVSNAGSITVPTIILVLDLRPGPGPI 18997  
QY 245 ----- 244  
Db 18998 RIDEVSCDSITISWNPPEYDGGCOISNYIVEKETTSTTHIVSOAVARTSIKIVRLTTG 19057  
QY 245 ----- 244  
Db 19058 SEYOFRVCAENRYCKSSYSESSAVVAEYFSPGPGTGVVHATKSTMLVTHQVPVNDG 19117  
QY 245 ----- 244  
Db 19118 GSRVIGYHLEYKERSILLWKNKILIAQTQVKVSGLDEGLMEYRVVAENIAGIGKCSK 19177  
QY 245 ----- 244  
Db 19178 SCEVPARDPCDPPGQPEVTNITRKSLSKWSKPHYDGGAKITGYIVERRELDPGRMLKC 19237



QY	245	-----	244
Db	19238	NYTNIQETVFEVTELTEDQRYEPRVFARNAADSVSEPSBSTGPIIVKDDVEPRVMDVK	19297
QY	245	-----	244
Db	19298	FRDVIIVKAGEVLKINADIAGRPLPVISWAKOGIEIERARTEIISTDNHLLTVKDCIR	19357
QY	245	-----	244
Db	19358	RDTGQYVLLKNVAGTRSAVNCKVLDKPPAGPLEINGLPAEKCSLSWGRQEDGAD	19417
QY	245	-----	244
Db	19418	IDYHRRKKRETSLAWTICEGELQMTCKVTLLKKGNEYIFRVTGVNKYGVGEPLSVAI	19477
QY	245	-----	244
Db	19478	KALDPFTVPSPPTSLEITSVTKESMTLCWSRPESDGGSEISGYIIBERREKNSLRWRVVK	19537
QY	245	-----	244
Db	19538	KPYDLRVKSTGLREGCEYEYRVYAENAGLSLPSSETPLIRAEDPVFLPSPSPKPIVD	19597
QY	245	-----	244
Db	19598	SGKTTITIAWVKPLFDGAPITGYTVEYKKSDTDWKTSIQSLRGTEYTIISGLTTGAEXV	19657
QY	245	-----	244
Db	19658	FRVKSVMKVGASDPSSDPQIAKEREEREPFLDIDSEMRKTLIVKAGASFTMTVPRGRP	19717
QY	245	-----	244
Db	19718	VPNVLSKPDTLTRAYVDTDSRTSLTIENANRDSKTYTLTQNVLSAASLTIVVKV	19777
QY	245	-----	244
Db	19778	LDTPGPTNITVQDVTKESAVLSWDVPENDGAPVKNYHIEKREASKAWSVTNCNRL	19837
QY	245	-----	244
Db	19838	SYKVTNLQEGAIYFRVSGENEFVGIPAEKTEGVKITEKPSPEKLGVTISIKDSVSLT	19897
QY	245	-----	244
Db	19898	WLKPEHGGSRIVHYVVALEKQKNWKCAYAKSTHHVVSGLRENSEYFFRVFAENQAG	19957
QY	245	-----	244
Db	19958	LSDPRELLLPVLKEQLEPPEIDMKNFPSHTVYVRAGSNLKVDPISGKPLPKVTLSDG	20017
QY	245	-----	244
Db	20018	VPLKATMRENTETAENLTINLKESVTADAGRYEITAANSSGTTKAFINIVLDRPGPPT	20077
QY	245	-----	244
Db	20078	GPVVISDITEESVTLKWEPPKYDGGSOVTNYILLKRETSTAVWTEYSATVARTMKVMKL	20137
QY	245	-----	244
Db	20138	TTGEEYQFRIKAENREGISDHIDSACVTVKLPYTPPGPSTPWNTVNTRESITVGNHEPV	20197
QY	245	-----	244
Db	20198	SNGGSVAVGYHLEMKDRNSILMQANKLVIRTHFKVTTISAGLIYEFRRYAENAAVGK	20257
QY	245	-----	247
Db	20258	PSHPSEPVLAIDACEPRNVRITDISKNSVLSWQQPAFDGGSKITGYIVERRDLPDGRW	20317
QY	248	-----	247
Db	20318	TKASFTNVTTOTFTISGLTQNSOYEFVRFARNAVGSISNPSEVYVGPITCIDSYGGPVIDL	20377
QY	248	-----	247
Db	20378	PLEYTEVWKYRAGTSVKLRAGISGKPAPTIETWKDDKELOTNALVCVENTTDLASILIKD	20437
QY	248	-----	247
Db	20438	ADRLNSGCYELKLARNAMASASATIRVOILDKPGPGPIEFKVTAEKITLLWRPPADDG	20497
QY	248	-----	247
Db	20498	GAKITHYIIVEKRETSRVWMSWSEHLEECIITTTIKIGNEYIFRVRVAVNKYIGEPLES	20557
QY	248	-----	247
Db	20558	DSVVAKNAFVTPGPGPIPEVTKITKNSMTVWNSRPIDAGSDISGYFLEKRDKSLGWFK	20617
QY	248	-----	247
Db	20618	VLKETIRDTRQKVTGLTENSDDYQYRVCANNAAGOGPSEPSEFYKAADPIDPPGPAKIR	20677
QY	248	-----	247
Db	20678	IADSTKSSITLWWSKPYDGSVATGYVVEIROGEEETWTVSTKGEVRTTEYVWSNLKP	20737
QY	248	-----	247
Db	20738	GVNYIFRVSANVACAGOGEPTEMNEPVQAKDILEAPEILDVALRTSVIAKAGEDVQLIP	20797
QY	248	-----	247
Db	20798	FKGRPPPTVTRKDEKNLGSADARYSIENTDSSLLTIPQVTRNDTGKYILTIENGVEPK	20857
QY	248	-----	247
Db	20858	SSTVSVKVLDTPAACOKLOVKHVSRTVTLLMDPPLIDGGSPINIVYIEKRDATKRTWSV	20917
QY	248	-----	247
Db	20918	VSHKCSSTSFKLIDLSEKTPFFFRVLAENEIGICEPCTTEPVKAAEVPAPIRDLMSKDS	20977
QY	248	-----	247
Db	20978	TKTSVILSMTKPDFDGGSVITEYVVERKKGEGQWTHAGISKTEIEVSQLEKQSVLEFR	21037
QY	248	-----	247
Db	21038	VFAKNEKGLSDPVTIGPITVKELIITPEVDLSDIPGAQVTRIGHNVHLELPYKGRPKPS	21097
QY	248	-----	247
Db	21098	ISWLKDGPLKESEFVRFSKTENKITLSIKNAKKEHGGKTYTILDNAVCRIAPITVITL	21157
QY	248	-----	247
Db	21158	GPSPKPKPIRDEIKADSVILSWDVPDNGGGEITCYSIEKRETSQTNWKMVCSSVART	21217
QY	248	-----	247
Db	21218	TFKVPNLVKDAEYQFRVRAENRYGVSOPLVSSIIIVAKHOPRIPGPPKPVYINWTS DGMS	21277
QY	248	-----	247
Db	21278	LTWDAVPYDGGSEVTFHVEKKERNILMQKVNTSPISGREYRATGLVEGLDYQFRVYAE	21337
QY	248	-----	247
Db	21338	NSAGLSSPSPSKFTLAVSPVDPPTDYIDVTREITLKNWPNPLDRDGGSKIYGVYSIEKR	21397
QY	248	-----	247

Db 21398 QGNERWVRNFTDVSECOYTVTGLSPGDRYEFRIIARNAVGTISPPSQSGIIMTRDENV 21457  
QY 248 ----- 247  
Db 21458 PPIVEFGPEYFDGLIIKSGESURIKALVOGRVPRVTFKDGVEIEKRMNMEITNVLGST 21517  
QY 248 ----- 247  
Db 21518 SLFVRDATRDHRGVYTVAKNAGSAAKAEIKVKVQDTPGKVVGPIRFTNITGEKMTLWWD 21577  
QY 248 ----- 247  
Db 21578 APLNDGAPITHYIIIEKRETSRLAWALIEDKCEAQSYTAIKLINGNEYQFRVSAVNKFCV 21637  
QY 248 ----- 247  
Db 21638 GRPLSDPVAQIQYVDPAPGIPSPNITGNSITLTWARPESDGGSEIQOYILERREKK 21697  
QY 248 ----- 247  
Db 21698 STRWVKVISKRPISETFRFKVTGLTEGNEYEFHVAENAGVGPAISRLIKCREPVNPP 21757  
QY 248 ----- 247  
Db 21758 GPPTVVVKVTDTSKTVSLEWSKVPFDGGMIEIIGYIIEMCKTDLGDWHKVNABEACVKTRYT 21817  
QY 248 ----- 247  
Db 21818 VTDLQAGEYKFRVSAINGAGKDCSEVTGTTIKAVDRLTAPELDIDANFKQTHVVRAGAS 21877  
QY 248 ----- 247  
Db 21878 IRLFIAYOGRPTTAVWSKPDNSLSRADIHDTDSFSTLTVCNCRNDAGKYTLAVENNS 21937  
QY 248 ----- 247  
Db 21938 GSKSITFTVKVLDTPGPPGPIITFKDVRGSATLMDAPLLDGGARIHHYVVEKREASRRS 21997  
QY 248 ----- 247  
Db 21998 WQVISEKCTRQIFKYNDLAEGVPYFRVSAVNEYGVGPEYEMPEIVATEQPAPRRLDV 22057  
QY 248 ----- 247  
Db 22058 VDTSSAVLAWLKPDPHDGSRITCYLLEMROKGSDLWVEAGHTKOLFTTVERLVEKTEY 22117  
QY 248 ----- 247  
Db 22118 EFRVAKNDAGYSEPREAFSSVIIKEPQIEPTADLTGITNOLITCKAGSPFTIDVPISGR 22177  
QY 248 ----- 247  
Db 22178 PAPKVTWKEENRLKETDRVSITTTIKDRTTLTVKDSMRGDSGRYPLTLENTAGVKTFSVT 22237  
QY 248 ----- 247  
Db 22238 VVVIGRPGVPTGPIEVSSVSAESCVLWGPCKDGGTEITNYIVEKRESGTTAWQLVNSS- 22297  
QY 248 ----- 247  
Db 22298 VKRTOIKVTHLTKYMEYSFRVSSNRFVGSKPLESAPIAIEHPFVPPSAPTRPEVYHVSA 22357  
QY 248 ----- 247  
Db 22358 NAMSIRWEEPYHDGSKIIGYVVEKKERTILWVKENKVPCLCNYKVTGLVEGLEYOFR 22417  
QY 248 ----- 247  
Db 22418 IYALNAGYSKASEASRPIMAQNVPDAPGRPEVTDVTRSTVLSIWSAPAYDGGSKVVGVI 22477  
QY 248 ----- 247  
Db 22478 IERKPVSEVGGRWLKCNVTIVSDNFTVTALSEGDTYEFVRVLAKNAAGVTSKSESGTP 22537

QY 248 ----- 247  
Db 22538 VTCRDEYAPPKAELDARLHGLVTTIRAGSDLVLDAAVGGKPEPKIITWTKGKELDCEKV 22597  
QY 248 ----- 247  
Db 22598 SLOYTGKRATAVIKFCDRSDSGKYTLTVKNASGTRKAVSMVKVLDSPGCGKLTYSRVTO 22657  
QY 248 ----- 247  
Db 22658 EKCTLAWSLPQDGAETHYIVERRETSLNWNIVEGECPTLSYVVVTRLIKNNEYIFRV 22717  
QY 248 ----- 247  
Db 22718 RAVNKYGPVVESEPIVARNSTFTIPSPGPIEEVGTGKEHIIOWTKPESDGGNEISNY 22777  
QY 248 ----- 247  
Db 22778 LVDRKRESLRWRVYNKYVYVYDTRLKVTSLMEGCDYQFRVTAVNAAGNSEPSESRNFTS 22837  
QY 248 ----- 254  
Db 22838 CREPSYTPGPPSAPRVVDTTKHSISLAWTKPMYDGGTDIVGVLEMQEKDQWYRVHTN 22897  
QY 255 ----- 254  
Db 22898 ATIRNTEFTVPLKMGOKYSFRVAANVVKMGHSEYSESTAEIEPVERIERIPDLELADLLK 22957  
QY 255 ----- 254  
Db 22958 TWTIRAGASRLWVSVSGRPPVITWSKQIDLASRAIIDTTESYSLLIIVKVNRYDAGK 23017  
QY 255 ----- 254  
Db 23018 YTIEAENOSGKSAFVLVKVYDTPGPCPSVKVKEVSRDSVTITWEIPTIDGGAPINNYIV 23077  
QY 255 ----- 254  
Db 23078 EKREAAARAFKTVTKCKSLYRISGLVEGTMHYFRVLPENYIGIGEPCESTDVAVLSEV 23137  
QY 255 ----- 254  
Db 23138 PLVPAKLEWVDVTKSTVTLAWEKPLYDGGSLTGYVLEACKAGTRMVKVVTLPKPTVLEH 23197  
QY 255 ----- 254  
Db 23198 TVTSLNEGEQYLFIRIRAONEKGVSEPRETVTAVTVQDLRVLPITDLSMPQKTIHVPAGR 23257  
QY 255 ----- 254  
Db 23258 PVELVIPIAGRPPPAASHFFAGSKLRESERVTVETHTKAKLITRETTIRDTGTYTLEK 23317  
QY 255 ----- 254  
Db 23318 NVTGTTSETIKVIIIDKPGPTGIKIDEIDATISITISWEPPELDGGAPLSGYVVEQRDA 23377  
QY 255 ----- 254  
Db 23378 HRPGLPVSESVTRSTFKFTRLTEGNEYVFRVAATNREFGYSYLOSEVIECRSSIRIPGP 23437  
QY 255 ----- 254  
Db 23438 PETLIQIFDVSRDGMTLWYPPEDDGGSOVTGYIVERKEVRADRWVRVNVKVPVMTTRYST 23497  
QY 255 ----- 254  
Db 23498 GLTEGLEYEHRVTAIARGSGKSRPSKPIVAMDP IAPPGKQPNPRVTDTRTTSVSLAWS 23557  
QY 255 ----- 254  
Db 23558 VPEDEGGSKVTVYLIEMQKVDQHEWTKONTPTTKIREYTLTHLPOGAERYFRVLACNAGG 23617

QY 255 -----KVAEKL----- 260  
|||  
Db 23618 PGEPAEVPGTGVKTEMLEYDPYELDERYQEGIFVRQGVIRLTIPKGPCKWTKEG 23677  
QY 261 ----- 260  
Db 23678 QDISKRAMIATSETHTELVIKEADRGSGTVLDVLENKCGKAVYIKVRVIGSPNPEGP 23737  
QY 261 ----- 260  
Db 23738 LEYDDIQVRVSWRPADGGADILGYILERREVPKAAWTTIDSRVGTSLVVKGLKE 23797  
QY 261 ----- 260  
Db 23798 NVEYHFRVSAENQFGISKPLKSEEPVTKPLNPPEPPNPPELDVTKSSVLSWSRPK 23857  
QY 261 ----- 260  
Db 23858 DGGSRVTGYIERKETSTDKVVRHNTQITTTMYTGLVPDAEYQFRIIAQNDVGLSE 23917  
QY 261 ----- 260  
Db 23918 TSPASEPVCKDPFKPSQPGCELEILSISKDSVTLOWEKPECDGKELGYWVEYRQSGD 23977  
QY 261 ----- 260  
Db 23978 SAWKSNKERIKDKQFTIGGLEATEYEFVFAENETGLSRPRRTAMSIKTKLTSGEAPG 24037  
QY 261 ----- 260  
Db 24038 IRKEMKDVTKLGEAQLSCQIVGRPLPDIKWYRFGKELIQSRKYMSSDGRTHTLVMT 24097  
QY 261 ----- 260  
Db 24098 EEQDEGYTCIATNEVGEVETSSKLLQATPQHPGYPKKEYYAGVSGTLRLHVMVIG 24157  
QY 261 ----- 260  
Db 24158 RVPAMTWFGKLLQNSENITIENTHYTHLVNMKNQKTHAGKYKQLSNVFGTVDAI 24217  
QY 261 ----- 260  
Db 24218 LDVEIQDKPDKPTGPIVIEALLKNSAVISWKPADGGSWITNYVVEKCEKAEGWQLV 24277  
QY 261 ----- 260  
Db 24278 SSAISVTCRIVNLTENAGYFRVSAQNTFGISDPLEVSSVVIKSPFEKPGAPKPTIT 24337  
QY 261 ----- 260  
Db 24338 AVTKDSCVWANKPPASDGGAKIRNYLEKREKKONKWSVTTEIRETVFSVKNLIEGLE 24397  
QY 261 ----- 260  
Db 24398 YEFVKCNLGESEWSEIPEITPKSDVP IQAPHFKELNLNVYQSNATLVCKVTGH 24457  
QY 261 ----- 260  
Db 24458 PKPIVKNYRQCKEIIADGLKYRIOEFKGYHQLIIASVTDDATVYQVRATNGSVSGT 24517  
QY 261 ----- 260  
Db 24518 ASLEVEVPAKTHLPKLTLEGMGAVHALRGEVVSIIKIPESGKPDVITWQGDLDINNGHY 24577  
QY 261 ----- 260  
Db 24578 QVIVTRFTSLVPNGVERKPDAGYVVCANKNRFGIDQKTVELDVADVPDPGRVKVSDAS 24637  
QY 261 ----- 260  
Db 24638 RDSVNLWTPEASDGGSKITNYIIVEKCATTAERWLRYQARETRYVINLFGKTSYQFRV 24697  
QY 261 ----- 260

Db 24698 IAENKFGLSKPSPESEPTITREKTRAMNYDEVDRETSWKASHSTKELYKYMIA 24757  
QY 261 ----- 260  
Db 24758 EDLGRGEGIVHRCVETSSKTTYMAKFKVKGTDQVLVKKEISILNIARHRNHLHESF 24817  
QY 261 ----- 260  
Db 24818 ESMEELWIFFEISGLDIFERINTSAFELNEREIVSYHVQVCEALQFLHSHNIGHFDIRP 24877  
QY 261 ----- 260  
Db 24878 ENIIYQTRRSSTIKIIEFGARQLKPGDNFRLLFTAPEYAYEVEHQHDVWSTATDWSLIG 24937  
QY 261 -----KEAF 264  
:|||  
Db 24938 TLVVYLLSGINPFLAETNQIENIMNAEYTFDEEAF 24974  
  
RESULT 15  
Q8WZB3  
ID Q8WZB3 PRELIMINARY; PRT: 26926 AA.  
AC Q8WZB3;  
DT 01-MAR-2002 (Tremblrel. 20, Created)  
DT 01-MAR-2002 (Tremblrel. 20, Last sequence update)  
DT 01-JUN-2002 (Tremblrel. 21, Last annotation update)  
DE N2B-titin isoform.  
GN TTN.  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
OX NCBI\_TaxID=9606;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=20309627; PubMed=10850961;  
RA Freiburg A., Trombitas K., Hell W., Cazorla O., Fougereousse F.,  
RA Centner T., Kolmerer B., Witt C., Beckmann J.S., Gregorio C.C.,  
RA Granzier H., Labeit S.;  
RT "Series of exon-skipping events in the elastic spring region of titin  
as the structural basis for myofibrillar elastic diversity";  
RL Circ. Res. 86:1114-1121(2000).  
RN [2]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=21573839; PubMed=11717165;  
RA Bang M.L., Centner T., Fornoff F., Geach A.J., Gotthardt M.,  
RA McNabb M., Witt C.C., Labeit D., Gregorio C.C., Granzier H.,  
RA Labeit S.;  
RT "The complete gene sequence of titin, expression of an unusual -700  
kDa titin isoform and its interaction with obscurin identify a novel  
Z-line to I-band linking system";  
RL Circ. Res. 89:1065-1072(2001).  
DR EMBL; AJ277892; CAD12455.1;  
DR InterPro; IPR000282; Cytok\_receptor\_2.  
DR InterPro; IPR000719; Euk\_pkinase.  
DR InterPro; IPR000577; FGKY\_kin.  
DR InterPro; IPR003961; FN\_III.  
DR InterPro; IPR001092; HLH\_basic.  
DR InterPro; IPR003599; Ig.  
DR InterPro; IPR003598; Ig\_c2.  
DR InterPro; IPR003600; Ig\_like.  
DR InterPro; IPR003006; Ig\_MHC.  
DR InterPro; IPR002016; Peroxidase.  
DR InterPro; IPR002290; Ser\_thr\_pkinase.  
DR InterPro; IPR001245; Tyr\_pkinase.  
DR Pfam; PF000041; fn3; 132.  
DR Pfam; PF000047; ig; 91.  
DR Pfam; PF000069; pkinase; 1.  
DR ProDom; PD000001; Euk\_pkinase; 1.  
DR SMART; SM00060; FN3; 132.  
DR SMART; SM00409; IG; 111.  
DR SMART; SM00408; IGC2; 93.  
DR SMART; SM00410; IG\_like; 3.



QY 43 -----VVVDVSHD----- 51  
Db 2078 IYWWPEDNVCLEVRDVTAEASIMVKAINIAGETSSHAFLVQAKQLITTEQLQDV 2137  
QY 52 -----PEVKFNWYDGVVH----- 66  
Db 2138 VAKEKDTMATFECETSEPFVKVWKYKMGMEVHEGDKYRMHSDRKVHFLSILITDSDAED 2197  
QY 67 -----NAKT----- 70  
Db 2198 YSCVLVEDENVKTAKLIVEGAVVEFVKELQDIEVPSYSGELECIVSPENIEGKWYHND 2257  
QY 71 -----KPR----- 73  
Db 2258 VELKSNGKYTITSRRGRQNLTVKDVTKEDQGEYSFVIDGKTKCKLMKPRPAILQGLS 2317  
QY 74 ----- 73  
Db 2318 DQKVCSEGDIVOLEVKVLSVESVEGWMKGOEQVQPSDRVHIVIDKQSHMLLIEDMTKEDAG 2377  
QY 74 ----- 73  
Db 2378 NYSFTIPALGSLTSGRVSYSVDVITPLKDVNVIEGTKAVLECKVSPDVTSVKWLINDE 2437  
QY 74 ----- 73  
Db 2438 QIKPDDRQVAIVKGTOKORLVINRTHASDEGPYKLIIVGRVETNCNLSVEKIKIIRGLRDLT 2497  
QY 74 ----- 73  
Db 2498 CTETQNVFEVLSHGIDVLNMFNFKDKEIKPSSKYKIEAHGIYKLTVLNMMKDDEGKYT 2557  
QY 74 ----- 73  
Db 2558 FYAGENITSGKLTVAGSAISKPLTDQTVABEQAEVPECEVANPDCKGEWLROGKHLPLTN 2617  
QY 74 ----- 73  
Db 2618 NIRSESDGHKRLIIAATKLDIDIGEYTKVATSKTSAKLKVEAVKIKTLKNTLVTETQD 2677  
QY 74 ----- 73  
Db 2678 AVFTVELTHPNVGVQWIKNGVNVLESNEYAISVKGTIYSLRIKNCIAVDESUYGFRLGR 2737  
QY 74 ----- 73  
Db 2738 LGASARLHVETVKIIKKPKDVTALENATVAFESVSHDTPVVKWFHKNVEIKPDKHRLV 2797  
QY 74 ----- 73  
Db 2798 SERKVHKLMLQNLSPSDAGEYTAVGQLECKAKLFFVETLHITKTMKNIEVPETKTASFEC 2857  
QY 74 ----- 73  
Db 2858 EVSHFNWPSMLKNGVEIEMSEKFIWVQCKLHQLIIMNTSEDSAEYTFVCGNDQVSAT 2917  
QY 74 ----- 73  
Db 2918 LTVTPIMITSMKLDINAEKDTITFEVTNYEGISYKWLKNGVEIKSTDKCMQRTKKLTH 2977  
QY 74 ----- 73  
Db 2978 SLNIRNVHFGDAADYTFVAGKATSTATLYVEARHIEFRKHIDIKVLEKKRAMFECEVSE 3037  
QY 74 ----- 73  
Db 3038 PDITVOMKDDQELQITDRIKIOKEYVHRLIPLIPSTRMSDAGKYTVVAGGNVSTAKLFVE 3097  
QY 74 ----- 73  
Db 3098 GRDVRIRSIKKEVQVIEKQRAVVEFVNEEDDDAHWKDGIENFQVQERHKYVVERRIH 3157  
QY 74 ----- 73

Db 3158 RMFISETRQSDAGEYTFVAGNRSSVTLTYVNAPEPPQVLOELQVTVQSGKPARCAVIS 3217  
QY 74 -----EE----- 75  
Db 3218 GRPQPKISWKEBQELLSTGFKCKFLHDGQBYTLLLIBAFPEDAAVYTCEAKNDYGVATTS 3277  
QY 76 ----- 75  
Db 3278 ASLSVEVEPVVSDQEMPVYPVPAIITPLQDVTVTVSEGQPARFQCRVSGTDLKVSHYSKDKK 3337  
QY 76 -----QYNSTYRV----- 83  
Db 3338 IKPSRFRMTQFEDTYQLEIAEAYPEDEGTYTFVASNAGQVSVSTANLSLEAPESILHER 3397  
QY 84 ----- 83  
Db 3398 IEQIEIEMEMKEFSSFLSABEEGLHSAELQSKINETLELLSESPVYTKFDSKEGTGP 3457  
QY 84 -----VSVL----- 87  
Db 3458 IFIKEYSNADISMGDVATLSVTYIGIPKIQHFFNGVLLTPSADYKFEVFDGDDHSLIIL 3517  
QY 88 ----- 87  
Db 3518 FTKLEDEGEYTCMASNDYGTICSAYLKINSKGEHGDTESETSAVAKSLEKLGCPPHF 3577  
QY 88 -----TVLH----- 91  
Db 3578 LKELKPIRCAOGLPAIFEYTVWGEPAFTVTFWKENKOLCTSVYVTIITHNPNGSTFIVND 3637  
QY 92 ----- 91  
Db 3638 PQREDSGLYICKAENMLGESTCAEALLVLEDDTMDTDTCKAKSTPEAPEDFPOTPLKGP 3697  
QY 92 ----- 91  
Db 3698 AVEALDSEQEIATFVKDTILKAALITEENQOLSIEHIAKANELSSQLPLGAQELQSILEQ 3757  
QY 92 ----- 91  
Db 3758 DKLTPESTREFLCINGSIHFOPLKEPSNLOLQIVOSQKTFSEKGIILMPEPEPTQAVLSD 3817  
QY 92 ----- 91  
Db 3818 TEXIFPSAMSIQINSITVEPLKTLAEPEGNYPQSSIEPPMHSYLTSVAEEVLSPEKT 3877  
QY 92 ----- 91  
Db 3878 VSDTNREQRTLOKQEAQSALILSQAEGHVESLQSPVVMISQVNYEPLVPSEHSCTEG 3937  
QY 92 ----- 91  
Db 3938 GKILIESANPLENAGQDSAVRIEGBKSLRPLALEEKQVLLKEEHSNVVMPDQIIESK 3997  
QY 92 -----QDWL- 95  
Db 3998 REPVAIKKVOGVRDLSKESLLSGIPEQRNLKTIQICRALQAAVASBQPLFSEWLR 4057  
QY 96 ----- 95  
Db 4058 NIEKVEEAVNITQEPRHIMCYLVTSAKSVTEEVTTIIEDVDPMANLKMELDALCAI 4117  
QY 96 ----- 95  
Db 4118 IYEBIDILTAEGPRIQOGAKTSLQEMDSFGSKQKVEPIPEPEVESKYLISPEEVSYFNV 4177  
QY 96 ----- 95  
Db 4178 QSRVKYLDATPVTKGVASAVVDEKQDESLLKPSKEESESSESTEVEATVKIQEABGGF 4237  
QY 96 ----- 95

Db 4238 IKEDGPMIHTPLVDTVSEGDIVHLLTTSITNAKEVNWYPENKLVPSDEKFKCLQDQNTYT 4297  
QY 96 -----NGK----- 98  
Db 4298 LVIDKVNTEHDHQGEYVCEALNDSGKTATSAKLTVVKRAAPVIRKIEPLEVALGHLAKFT 4357  
QY 99 -----EYKCKVSN----- 106  
Db 4358 CEIQSAPNVRFOWKAGREIYESDKSIRSSKYISSLEILRTQVVDCGEYTKCASNEYS 4417  
QY 107 ----- 106  
Db 4418 VSCTATLTVTVPGEKKVKRLLPERKPEKEEVLKSVLRKRPEREPEKVKLEKVKK 4477  
QY 107 ----- 106  
Db 4478 PAVPEPPPKPVEEVPVTRKRKIPEPTKVPEIKPAIPAPAPKPKPEAEVKTIKP 4537  
QY 107 -----KALPAPIE----- 114  
Db 4538 PPVEPEPTPIAAPVTVPVVGKKAEPKAAKPKGPIKGVPKKTPSPIEAERRKLKRP 4597  
QY 115 ----- 114  
Db 4598 SSGEKPPDEAPTYOLKAVPLKVFKEIKDIIILTESEFVGSSEFVCLVSPSTAITWMKD 4657  
QY 115 ----- 114  
Db 4658 GSNIRESPKRFIADGDKRKLHIIDVQLSDAGEYTCVLRGNKEKTSTAKLVVEELPVRF 4717  
QY 115 -KTIS----- 118  
Db 4718 VRTLEEVTVKQPLYLSCLENKERDVVRKDGKIWEKPGRIVPGVIGLMALTIINDA 4777  
QY 119 ----- 118  
Db 4778 DTDAGTYTTVVENANNLECCSVKVEVIRDWLKVPIRDQHVKPKGTAIFACDIADKTP 4837  
QY 119 ----- 118  
Db 4838 NIKWFGYDEIPAENDKTEILRDGNHLYLKIKNAMPEDIAEYAVEIEGKRYPAKLTGE 4897  
QY 119 ----- 118  
Db 4898 REVELLKPIEDVTIYEKESAFDAITSEADIPGQWKLKGLLSPSTCEIKAEKGKREIT 4957  
QY 119 ----- 118  
Db 4958 LRKVKLDQAGEVLYOALNAITTAITLVKEIFELDFAVPLKDVTPERRQARFECVLTREAN 5017  
QY 119 ----- 118  
Db 5018 VIWSKGPDIKSSDFDIADGKKHILVINDSQFDDGEVYTAIEVEGKTSARLFTVGTGRL 5077  
QY 119 ----- 118  
Db 5078 KPMSPLEQTVKEGATATFVCELSHEKMHVWFKNDAKLHSTRTVLISSEKGTHELMKE 5137  
QY 119 ----- 118  
Db 5138 VTLDDISQIAQVKELSSAQKLVLEADPYFTVKLHDKTAVEKDEITLKCEVSKDVPVKW 5197  
QY 119 ----- 118  
Db 5198 FKDGEIIVSPKYSIKADGLRILKIKADLKDKGEYVDCGTDKTKANVTVEARLIKVE 5257  
QY 119 -----KAGQP----- 124  
Db 5258 KPLYGVEFVGETAHEIELSEPVHGWKLKGLQPLTASPCDEIIEDGKKHILLHNCQL 5317  
QY 125 -----REPO----- 128  
Db 5318 GMTGEVSQAAAKSAANLKVKEPLIFITPLSDVKVFEKDEAKFECEVSREPKTFRWLK 5377

QY 129 ----- 128  
Db 5378 GTQEITGDDRFELIKDGTKHSMVIKSAAFEDEAKYFMEAEDKHTSGKLIIEGIRLKLFTP 5437  
QY 129 ----- 128  
Db 5438 LKDVTAKEKESAVFTVELSHDNIRVKFNQDORLHTTRSVMQDEGKHTSIITFKDLSIDD 5497  
QY 129 ----- 128  
Db 5498 TSQIRVEAMGMSSEAKLTVLEGDYFTGKLQDYTGVEKDEVILQCEISKADAPVKWFKDG 5557  
QY 129 ----- 128  
Db 5558 KEIKPSKNAVAKADGKRRMLILKALKSDIGQYTCDCGTDKTSGLDIEDREIKLVRPLH 5617  
QY 129 ----- 128  
Db 5618 SVEVMEETARPETEISEDDIHANNKLKGEALLQTPDCEIKEEGKIHSLSVLHNCRLDQGT 5677  
QY 129 ----- 128  
Db 5678 GVDFOAANVYKSAHLRVKPRVIGLLRPLKDVTVTAGETATFDCELSYEDIPVEWYLGKK 5737  
QY 129 ----- 128  
Db 5738 LEPSDKVVPSEGVHTLTLRDVKLEDAGEVOLTAKDFKTHANLFLVKEPPVEFTKPLEDQ 5797  
QY 129 ----- 128  
Db 5798 TVEEGATAVLECEVSRENKVKWFKNGTEILSKKYEIVADGRVRKLVITHDCTPEDIKTY 5857  
QY 129 ----- 128  
Db 5858 TCDAKDFTSCNLNVVPPHVEFLRPLTDLVREKEMARFECELSRENKVKWFKDGAETK 5917  
QY 129 ----- 128  
Db 5918 KGKDYDIISKGAVRILVINKCLLDDAEYSCVETARTSGMLTVLEEEAVFTKNLANIEV 5977  
QY 129 ----- 128  
Db 5978 SETDTIKLVCEVSKPAEVIWYKGDDEIETGRYEILTEGRKRIVLVIONAHLEDAGNYC 6037  
QY 129 ----- 128  
Db 6038 RLPSSRTDGKVKVHELAAEFISKPONLEILEGKAEFVCSISKESFPQMKRDDKTLESG 6097  
QY 129 ----- 128  
Db 6098 DKYDVIADGKKRVLVVKDATLQDMGTYYVMVGAARAAHLTVIEKLRIWVPLKDTVRKEQ 6157  
QY 129 -----VYTL----- 132  
Db 6158 QEVFVCEVNTGAKAKWFRNEEAFDSSKYILQKDLVYTLRIRDAHLDDQOANYSLT 6217  
QY 133 ----- 132  
Db 6218 NHRGENVSAANLIVEEEDLRIVEPLKDIETMEKKSFTWCKVKNRLNVTLWTKNGEVP 6277  
QY 133 ----- 132  
Db 6278 FDNRSYRVKYYKHLMTIKDCGFPDEGEYIVTAGODKSVAEILLIIEAPTEFVEHLEDQTV 6337  
QY 133 ----- 132  
Db 6338 TEFDDAVFSCQLSREKANVKWYRNGREIKGKFKFKDGSIHRLI IKCDRLDDECEYAC 6397  
QY 133 ----- 132  
Db 6398 GVEDRKSARLFEVEIPVEIIRPPQDILEAPGADVVFVLAELNKKVQWLNNMVVQVG 6457

QY	133	-----	132
Db	6458	DKHQMSEGIHRLQICDIPROGEYRFTAKDEARAKLELAAAPKIKTADQDLVDVG	6517
QY	133	-----	132
Db	6518	KPLTMVVPYDAYPKAAEFKENEPLSTKTIDTTAEQTSFRILEAKKGDKRYKIVLQNK	6577
QY	133	-----	132
Db	6578	HGAEGFINLKVIDVPGVVRNLEVTFDFGEVSLAMEEPLTDGGSKIIGYVVERDIKRK	6637
QY	133	-----	132
Db	6638	TWVLATDRAESCEFTVTGLQGGVEYLFVRSARNRVGTGPEVDNPNVEARSKYDVGPP	6697
QY	133	-----	132
Db	6698	LNVITIDVNRFGVSLWEPPEYDGAETINYVIELRDKTSIRWDTAMTVRAEDLSATVTD	6757
QY	133	-----	132
Db	6758	VVEGEYSFRVRAQNRIGVCKPSAATPFVKVADPIERPSPVNLTSDDQTQSSVOLKWEF	6817
QY	133	-----	132
Db	6818	PLKGGSPILGYIIERCEECKDNWIRCNMKLVPELTYKVTGLEKGNKLYRVSANKAGV	6877
QY	133	-----	132
Db	6878	SDPSEILGPLTADDAFVEPTMDLSAFKDGLEIVPNPITILVPSTGYPRPTATWCFQDKV	6937
QY	133	-----	132
Db	6938	LETGDRVKMKTLSAYAEVLSPSERSDKGIYTLKLENRVKTSIGEDVNVVIARSPAPKEL	6997
QY	133	-----	132
Db	6998	KFGDITKDSVHLATWEPDDGGSLTGYVVEKREVSRTKWTVMDFVTDLEFTVPLVQG	7057
QY	133	-----	132
Db	7058	KEYLFKVCARNKCGPEPAYVDEPNMSTPATVDPDPENVKWRDRTANSIFLTWDPKND	7117
QY	133	-----	132
Db	7118	GGSRIGYIVERCPGSKWVACGEPVAETKMEVTGLEGKWTAYRVKALNRQASKPSR	7177
QY	133	-----	132
Db	7178	PTEEIOAVDTQEAPEIFLDVKLLAGLTAKGTRKIELPATVTGKPEKITWTKADMILKQD	7237
QY	133	-----	132
Db	7238	KRITIENVPKSTVTIVDSKRSRTGYIIEAVNVCGRATAVEVNVLDKPGPAAFDITD	7297
QY	133	-----	132
Db	7298	VTNESCLLTWNPPRDDGSKITNYVVERRATDSEVHKLSTVTKDTNFKATKLIPNKEYI	7357
QY	133	-----	147
Db	7358	FRVAENMYGVEPVQASPIKAYQFDPGPPTRLEPSDITKDAVLTWCEPDDGGSPI	7417
QY	148	-----	147
Db	7418	TGYVVERLDPDQKWRNCNMPVKDITYRVKGLTNKKYFRVLAENLAGPGKPSKSTEP	7477
QY	148	-----	147
Db	7478	ILIKDIPDPWPGKPTVKDVGKTSVRLNWKPEHGGAKIESYVIEMLKTGTDEWVRVA	7537
QY	148	-----	147

Db	7538	EGVPTTQHLPLGLMEGOEYSFRVRAVNKAGESESEPSDPLVLCREKLYPPSPRWLEVIN	7597
QY	148	-----	147
Db	7598	ITKNTADLKWTVPKDGGSPIITNYIVEKRDVRRKGWQVDTTIVKDKTRCTVPLTEGSLYV	7657
QY	148	-----	147
Db	7658	FRVAENAIQSDYTEIETEDSVLAKDTFTTGPYPYALAVDVKRHVDLKWEPKNDGGRP	7717
QY	148	-----	147
Db	7718	IQRVIEKKERLGTGRWKAGKTAGDCNFRVTDVIEGTEVQFQVRAENAGVHPSEPT	7777
QY	148	-----	147
Db	7778	ILSIEDTSPSPPLDLHVTDAGRKHIAIAWKPEKNGGSPIIIGYHVEMCPVGTEKWMRV	7837
QY	148	-----	147
Db	7838	NSRPDKLKFVEGVVPDKEYVLRVAVNAIGVSESEISENVVAKDPDCKPTIDLETH	7897
QY	148	-----	147
Db	7898	DIIVIEGEKLSIPVPFRAVPVPTVSWHKDGKEVKASDRLTMKNDHISAHLEVPKSVRADA	7957
QY	148	-----	147
Db	7958	GIYTITLENKLSATASINVKVIGLPGPCKDIKASDITKSKCLTWEPPEFDGTPILHY	8017
QY	148	-----	147
Db	8018	VLERRAGRRTYIPVMSGENKLSWTVKDLIPNGEYFRVAVNKVGGGEVIELKNPVAQ	8077
QY	148	-----	147
Db	8078	DPKQPPDPVDVEVHNPTAEAMTITWKPLPYDGGSKIMGYIIEKIAKEBRKRCNEHLV	8137
QY	148	-----	147
Db	8138	PILTYTAKLEGEKEYQFVRAENAAAGISEPSRATPPTKAVDPIADPKVILRTSLEVKRG	8197
QY	148	-----	147
Db	8198	DEIATDASISGSPYPTITWIKDENVIYVEIEIKKRAAPLVRRRKEGEVEEPFVLPLTQRL	8257
QY	148	-----	147
Db	8258	SIDNSKKGESQLRVDRSLRPHGLYMIKVENDHGIAKAPCTSVSLDTPGPPINFVFEDIR	8317
QY	148	-----	147
Db	8318	KTSVLCKWEPPDDGSEIINYTELKKDKTKPDSEIIVTSTLRHCKYSVTKLJEGKEYL	8377
QY	148	-----	147
Db	8378	FRVRAENRFGPPCYCKVPLVAKDPFGPDAPDKPIVEDYTSNMLVKNWNEPRDNGSPIL	8437
QY	148	-----	147
Db	8438	GYWLEKREVNSTHWSRVNKSILLNALKANVDGLLEGLTYVFRVCAENAAAGPKFSPSPDK	8497
QY	148	-----	147
Db	8498	TAHDPISPGPPPIPRVTDTSSTTIELEWEPFANGGGEIVGYFVDKOLVGTNEWSRCTEK	8557
QY	148	-----	147
Db	8558	MIKQVTVKEIREGADYKLRVSAVNAAGEGPGTPTVVAEPQEPAPVELDVSKGGI	8617
QY	148	-----	147

Db 8618 QIMAGKTLRIPAVTGRPVPTKWTKEGELDKDRVIDNVGTSKSELIIKDALKRHDGRY 8677  
QY 148 ----- 147  
Db 8678 VITATNSCGSKFAAARVEVDVPGVLDLKPVTNRKMKLLNWSDEDDGGSEITGFIE 8737  
QY 148 ----- 147  
Db 8738 RKDAKHTWROPIETERSKCDITGLLEGOEYKFRVIAKNKFCGPPVEIGPILAVDPLGP 8797  
QY 148 ----- 147  
Db 8798 PTPERLTYTERTKSTITILDMKEPRSGSGPIQGYIIIEKRHRDKPDFERVNKRCLCTTSF 8857  
QY 148 ----- 147  
Db 8858 LVENLDEHOMYEFVRKAVNKEISEPSPLNVVIODEVPTIKRLSVRGDTIKVKAGE 8917  
QY 148 ----- 147  
Db 8918 PVHIPADVTGLPMPKIENKNETVIEKPTDALQITKEEVSERSEAKTELSIPKAVREDKGT 8977  
QY 148 ----- 147  
Db 8978 YVTASNRLGVSFRNVHVEYDRPSPPRNLAVTDIKAESCYLTWDAPLDNGGSEITHYVI 9037  
QY 148 ----- 147  
Db 9038 DKRDASRKAEWEVTNTAVERKYGIIWKLI PNGQYEFVRVAVNKYGISDECKSKVVIQD 9097  
QY 148 ----- 147  
Db 9098 PYRLPGPGKPKVARTKGSMLVSWTPTPLDNGSGPITGVWLEKREEGSPYNSRVSRAPIT 9157  
QY 148 ----- 147  
Db 9158 KVLKGVEFNVRLLEGVGYQFRAMA INAAIGIGPPSPDEPVAGDPIFPFGPPSPCEVK 9217  
QY 148 ----- 147  
Db 9218 DKTKSSISLGNKPPAKDGGSPKIGYIVEMQEBGTTDKRVNPDKLIITCECVVFNKEL 9277  
QY 148 -----CLV----- 150  
Db 9278 RKYRFRKAVNAGESEPSDTTGEIPATDIOEPEVFIDIGAQCCLVCKAGSOIRPAVI 9337  
QY 151 ----- 150  
Db 9338 KGRPTPKSSWEPDGAKKAMKDGVDHIDIPEDAQLETAENSVIIIECKRSHTGKYSITAK 9397  
QY 151 -----KGF----- 153  
Db 9398 NKAGQKTANCRVKVMDVPGPKDLKVS DITRGSCRLSNKMPDDDDGGDRIGYVIEKRIID 9457  
QY 154 -----YPSD- 157  
Db 9458 GKAWTKANVDCGSTTFVVPDLLEQQYEFVRVRAENRFGIPGVETIQRTTARDPIYPPD 9517  
QY 158 ----- 157  
Db 9518 PIKLIGLITKNTVHLSWKPKPKNDGGSPVTHYIVECLAWDPTGTTKEAWRQCNKRDEEL 9577  
QY 158 ----- 157  
Db 9578 QFTVEDLVEGGEYEFVRKAVNAAGVSKPSATVGPCDCQCPDMPSPSIDLKEFMEVEEGTNV 9637  
QY 158 ----- 157  
Db 9638 NIVAKIGVPEPTLTFWFKAPKPKONKEPVLVDTHVNLVVDCTCTLVIPQSRSDTGLY 9697  
QY 158 ----- 157  
Db 9698 TITAVNNLGASKEMRLNVLGRPGPPGPIKFSVSAQOMTLSWPPPKDDGGSKITNYVI 9757

QY 158 ----- 157  
Db 9758 EKREANRKTWVHSSEPKECTYTIKLLLEGHEYVPRIMAQNKYGICEPLDSEPETARNLF 9817  
QY 158 -----IAVEWE----- 163  
Db 9818 SVPGAPDKPTVSSVTRNSMTVNWEBPEYDGGSPVTGYWLEMKDTTSKRWKRVRNROPKAM 9877  
QY 164 ----- 163  
Db 9878 TLGVSYKVTGLIEGSDYQFRVYA INAAGVGPASLPSPATARDPIAPGPPFPKVTDMTK 9937  
QY 164 ----- 163  
Db 9938 SSADLENSPPLKDGSKVTGYIYEYKEEGKEWEKGDKEVRGRTLVTVTGLKEGAFYKFR 9997  
QY 164 ----- 163  
Db 9998 VSAVNIAGIGEPGEVTDVIEMKDRLVSPDLQDASVRDRIVVHAGGVIRIIAIVVSGKPPP 10057  
QY 164 ----- 163  
Db 10058 TVTWMNERTLPQEATIIETAISSVMVKNQCORSHQGVYSLAKNEAGERKKTIIIVDVL 10117  
QY 164 -----SNGOPENNY----- 172  
Db 10118 VPGPVGTPLAHLNLTNESCCLTWFSPEDDGGSPITNYVIEKRESORRAWTPVTYTVTRON 10177  
QY 173 ----- 172  
Db 10178 ATVQGLIQGKAYFFRIAAENSIGMPFVETSEALVIREPITVPERPEDLEVKEVTKNTVT 10237  
QY 173 ----- 172  
Db 10238 LTWNPVKYDGGSEIINYVLESRLIGTEKPHKVTNDNLLSRKYTVTKLKEGDTYEVRSVAV 10297  
QY 173 ----- 172  
Db 10298 NIVGOGKPSFCTKPTCKDELAPPTLHLDFRDKLTI RVGEAFALTRYSGKPKVSNPK 10357  
QY 173 -----KTP----- 176  
Db 10358 DEADVLEDDRTHIKTTPATLALAKIKAKRSDSGKYCVVVENSTGSRKGFQCVNVVDRPFP 10417  
QY 177 ----- 176  
Db 10418 PVGPVSFDEVTKDYMVISWKPPPLDDGGSKITNYIIEKKEVGKDWMPVTSASAKTTCKVS 10477  
QY 177 ----- 176  
Db 10478 KLLEGKDYIFRIHAENLYGISDPLVSDSMKAKDRFPVDPADPQPIVTEVTKDSALVTWKK 10537  
QY 177 ----- 176  
Db 10538 PHDGGKPIITNYILEKRETKSRWARVTKDPIHPYTKFRVPDLLEGCOYEFVRSAEINEIGI 10597  
QY 177 ----- 176  
Db 10598 GDPSPKPVFAKDIAPKPPVNPPEAIDTTCNSVDLTWQPPRHGGSKILGYIYEVQKV 10657  
QY 177 ----- 176  
Db 10658 GDEWRANRHTPESCPEYKVTGLRDQYKFRVLAVNAAGESDPAHVPEPVLVKDRLE 10717  
QY 177 ----- 176  
Db 10718 PPELILDANMAREQHIKVGDTLRLSAIIKGVFPKVTWKKEDRDAPTAKRIDVTPVSGKL 10777  
QY 177 ----- 176  
Db 10778 EIRNAAHEDGGIYSLTVENPAGSKTVSVKVLVLDKPGPPRDLEVSEIRKDCSYLTWKEPL 10837



GenCore version 5.1.4 p5 4578  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: April 21, 2003, 10:42:21 ; Search time 28.9412 Seconds  
(without alignments)  
1132.630 Million cell updates/sec

Title: 2LINK7  
Perfect score: 1334  
Sequence: 1 MDKTHRCPCPAPPELLGSPS.....GKDWLKAIFYDKVAEKLEAF 246

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.0

Searched: 908470 seqs, 133250620 residues  
Total number of hits satisfying chosen parameters: 908470

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database :

A.Geneseq.-101002:\*

- 1: /SIDS2/gcgcdata/geneseq/geneseq-emb1/AA1980.DAT:\*
- 2: /SIDS2/gcgcdata/geneseq/geneseq-emb1/AA1981.DAT:\*
- 3: /SIDS2/gcgcdata/geneseq/geneseq-emb1/AA1982.DAT:\*
- 4: /SIDS2/gcgcdata/geneseq/geneseq-emb1/AA1983.DAT:\*
- 5: /SIDS2/gcgcdata/geneseq/geneseq-emb1/AA1984.DAT:\*
- 6: /SIDS2/gcgcdata/geneseq/geneseq-emb1/AA1985.DAT:\*
- 7: /SIDS2/gcgcdata/geneseq/geneseq-emb1/AA1986.DAT:\*
- 8: /SIDS2/gcgcdata/geneseq/geneseq-emb1/AA1987.DAT:\*
- 9: /SIDS2/gcgcdata/geneseq/geneseq-emb1/AA1988.DAT:\*
- 10: /SIDS2/gcgcdata/geneseq/geneseq-emb1/AA1989.DAT:\*
- 11: /SIDS2/gcgcdata/geneseq/geneseq-emb1/AA1990.DAT:\*
- 12: /SIDS2/gcgcdata/geneseq/geneseq-emb1/AA1991.DAT:\*
- 13: /SIDS2/gcgcdata/geneseq/geneseq-emb1/AA1992.DAT:\*
- 14: /SIDS2/gcgcdata/geneseq/geneseq-emb1/AA1993.DAT:\*
- 15: /SIDS2/gcgcdata/geneseq/geneseq-emb1/AA1994.DAT:\*
- 16: /SIDS2/gcgcdata/geneseq/geneseq-emb1/AA1995.DAT:\*
- 17: /SIDS2/gcgcdata/geneseq/geneseq-emb1/AA1996.DAT:\*
- 18: /SIDS2/gcgcdata/geneseq/geneseq-emb1/AA1997.DAT:\*
- 19: /SIDS2/gcgcdata/geneseq/geneseq-emb1/AA1998.DAT:\*
- 20: /SIDS2/gcgcdata/geneseq/geneseq-emb1/AA1999.DAT:\*
- 21: /SIDS2/gcgcdata/geneseq/geneseq-emb1/AA2000.DAT:\*
- 22: /SIDS2/gcgcdata/geneseq/geneseq-emb1/AA2001.DAT:\*
- 23: /SIDS2/gcgcdata/geneseq/geneseq-emb1/AA2002.DAT:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1246	93.4	247	AA16958	FC-TMP protein seq
2	1246	93.4	247	AB173411	FC-TPO mimetic pep
3	1246	93.4	248	AB17953	FC-IL-1 antagonist
4	1246	93.4	248	AB173421	FC-interleukin 1 (
5	1246	93.4	268	AB16959	FC-TMP-TMP protein
6	1246	93.4	268	AB173412	FC-TMP-TMP amino a
7	1246	93.4	269	AA196531	Human IgG1 FC-TMP
8	1246	93.4	731	AA152156	Humanised HMG-1 h
9	1246	93.4	741	AA152159	Humanised HMG-1 h
10	1246	93.3	374	AA183963	Recombinant human

11	1245	93.3	374	AA19075	Recombinant human
12	1245	93.3	401	AA172922	Human met-Fc (lack
13	1245	93.3	401	AA180904	Human metFcDeltaC-
14	1245	93.3	423	AA186633	FC-huAaP-1 (114-28
15	1245	93.3	441	AA186692	FC-huAaP-1 (95-281
16	1245	93.3	689	AA196780	Ang-2-FD-Fc-FD fus
17	1245	93.3	729	AA152158	Humanised HMG-1 h
18	1245	93.3	739	AA152161	Humanised HMG-1 h
19	1244	93.3	652	AA18650	Heavy chain of hma
20	1243	93.2	633	AA184965	Amino acid sequenc
21	1242	93.1	252	AA179955	FC-VEGF antagonist
22	1242	93.1	252	AA173423	FC-VEGF antagonist
23	1242	93.1	689	AA196779	Ang-1-FD-Fc-FD fus
24	1241	93.0	426	AA186694	FC-huAaP-1 (120-29
25	1241	93.0	448	AA186694	FC-huAaP-1 (99-291
26	1240	93.0	379	AA183962	Recombinant human
27	1240	93.0	379	AA19073	Recombinant human
28	1240	93.0	379	AA19074	Recombinant human
29	1240	93.0	651	AA186649	Chimeric receptor
30	1240	93.0	692	AA186650	Chimeric receptor
31	1239	92.9	248	AA17951	FC-TNF-alpha inhib
32	1239	92.9	248	AA173419	FC-TNF-alpha inhib
33	1239	92.9	253	AA16964	FC-EMP protein seq
34	1239	92.9	253	AA173415	FC-EMP mimetic pep
35	1239	92.9	277	AA16967	FC-EMP-EMP protein
36	1239	92.9	277	AA173418	FC-EMP-EMP nucleic
37	1239	92.9	581	AA181972	Ganglioside GD2 sp
38	1239	92.9	582	AA181987	Ganglioside GD3 sp
39	1239	92.9	582	AA181991	Ganglioside GM2 an
40	1239	92.9	583	AA183156	Human IgG1 FC prot
41	1238	92.8	228	AA16955	Human IgG1 FC chai
42	1238	92.8	228	AA196523	Human IgG1 FC regl
43	1238	92.8	228	AA196953	Human IgG1 FC. Ho
44	1238	92.8	228	AA181074	Human immunoglobul
45	1238	92.8	228	AA173410	

#### ALIGNMENTS

RESULT 1	AA16958	AA16958 standard; Protein; 247 AA.
ID	AA16958	
AC	AA16958;	
DT	31-OCT-2000 (first entry)	
XX		
DE	FC-TMP protein sequence SEQ ID NO:6.	
XX		
KW	Modified peptide; therapeutic agent; fusion; Fc domain; cancer;	
KW	autoimmune disease; cytostatic; antitumour; thrombolytic; VEGF;	
KW	immunosuppressive; EPO; TPO; CTLA4; mimetic; IL-1; TNF; antagonist;	
KW	inhibitor; erythropoietin; thrombopoietin; interleukin 1;	
KW	cytotoxic T cell lymphocyte antigen 4; tumour necrosis factor;	
KW	vascular endothelial growth factor; matrix metalloproteinase;	
KW	asthma; thrombosis; pharmaceutical.	
XX		
OS	Homo sapiens.	
OS	Synthetic.	
PN	WO200024782-A2.	
XX		
PD	04-MAY-2000.	
XX		
PF	25-OCT-1999; 99WO-US25044.	
XX		
PR	23-OCT-1998; 98US-0105371.	
PR	22-OCT-1999; 99US-0428082.	
XX		
PA	(AMGE-) AMGEN INC.	
PI	Felge U, Liu C, Cheetham J, Boone TC;	

XX WPI: 2000-350702/30.  
 DR N-PSDB: AAA69444.  
 XX  
 PT Novel composition of matter comprising an FC domain and  
 PT pharmacologically active peptides, useful for treating cancer and  
 PT autoimmune diseases -  
 PS  
 PS Claim 21: Page 179-180; 608pp; English.  
 CC  
 CC The present invention describes composition of matter (I) comprising an  
 CC FC domain, pharmacologically active peptides, and linkers. Where (I) is:  
 CC (X1)a-E1-(X2)b, where: E1 = an FC domain; X1 and X2 = are each  
 CC independently selected from -(L1)c-P1, -(L1)c-P1-(L2)d-P2,  
 CC -(L1)c-P1-(L2)d-P2-(L3)e-P3, or -(L1)c-P1-(L2)d-P2-(L3)e-P3-(L4)f-P4  
 CC where P1, P2, P3, and P4 = are each independently sequences of  
 CC pharmacologically active peptides; L1, L2, L3, and L4 = are each  
 CC independently linkers; and a, b, c, d, e, and f = are each independently  
 CC 0 or 1, provided that at least 1 of a and b is 1. The composition can  
 CC have cytostatic, antitumour, thrombolytic and immunosuppressive  
 CC activities. DNAs, vectors and host cells from the present invention can  
 CC be used for producing pharmaceutical compositions. The compositions are  
 CC useful for treating cancer, asthma, thrombosis, or autoimmune diseases.  
 CC The use of an FC domain (rather than a Fab domain) can provide a longer  
 CC half-life or incorporate functions such as FC receptor binding, protein  
 CC A binding, complement fixation, and possibly placental transfer. AAA69443  
 CC to AAA69526 and AAB16955 to AAB18003 represent nucleotide and amino acid  
 CC sequences used in the exemplification of the present invention.  
 XX  
 XX Sequence 247 AA:  
 SQ  
 Query Match 93.4%; Score 1246; DB 21; Length 247;  
 Best Local Similarity 94.3%; Pred. No. 1.6e-35;  
 Matches 231; Conservative 0; Mismatches 2; Indels 12; Gaps 1;

QY 1 MDKTHCPCPAPPELLGGSVFLFPKPKDTLMISRTPEVTCVVDVSHEDPEVKFNWY 60  
 DB 1 MKDTHCPCPAPPELLGGSVFLFPKPKDTLMISRTPEVTCVVDVSHEDPEVKFNWY 60  
 OY 61 DGEVHNATKPREOYNSTYRVSVLVVLDHMDMLNGKREYKCKVSNKALPAIEKTISKA 120  
 DB 61 DGEVHNATKPREOYNSTYRVSVLVVLDHMDMLNGKREYKCKVSNKALPAIEKTISKA 120  
 OY 121 KGQPREPOVYTLPPSRDELTKNQVSLTCLVKGFYPSDIAVEMESNGQPENNYKTTPPVLD 180  
 DB 121 KGQPREPOVYTLPPSRDELTKNQVSLTCLVKGFYPSDIAVEMESNGQPENNYKTTPPVLD 180  
 OY 181 SGGSFELYSKLTVDKSRMOQGNVSCSVHMEALHNHYTQKSLSLSPGK-----228  
 DB 181 SGGSFELYSKLTVDKSRMOQGNVSCSVHMEALHNHYTQKSLSLSPGKGGGGIEGPTLR 240  
 OY 229 DWLKA 233  
 DB 241 QWLAA 245

RESULT 2  
 ABB73411  
 ID ABB73411 standard: Protein; 247 AA.  
 XX  
 XX ABB73411:  
 DE 05-APR-2002 (first entry)  
 XX  
 XX FC-TPO mimetic peptide (Fc-TMP) amino acid SEQ ID NO:6.  
 XX  
 XX Modified peptide; mimetic; Fc domain; fusion; immunoglobulin G; IgG;  
 KM EPO; erythropoietin; TPO; tumour necrosis factor alpha inhibitor;  
 KM TNF-alpha inhibitor; interleukin 1 antagonist; IL-1 antagonist; TMP;  
 KM TPO mimetic peptide; EPO mimetic peptide; EWP; VEGF antagonist;  
 KM MMP inhibitor; antinflammatory; antitumour; immunosuppressive;  
 KM cytostatic; antitumour; antitachycardic; antidiabetic; ophthalmological;  
 KM antinaemic; anorectic; antinfertility; haemostatic; dermatological;

KM neuroprotective; inflammatory disease; autoimmune disease; tumour growth;  
 KM cancer; rheumatoid arthritis; diabetic retinopathy; infertility; obesity;  
 KM sleep disorder; neurological degenerative disease; anaemia;  
 KM thrombocytopaenia; metastatic tumour; systemic lupus erythematosus;  
 KM Fanconi's syndrome.  
 XX  
 XX Homo sapiens.  
 OS  
 OS Synthetic.  
 XX  
 XX WO200183525-A2.  
 XX  
 XX 08-NOV-2001.  
 XX  
 XX 02-MAY-2001; 2001WO-US14310.  
 XX  
 XX 03-MAY-2000; 2000US-0563286.  
 PR  
 XX (AMGE-) AMGEN INC.  
 PA  
 PA Feige U, Liu C, Cheetham JC, Boone TC, Gudas JW;  
 PI  
 PI WPI: 2002-130313/17.  
 DR N-PSDB: ABL35761.  
 XX  
 XX Novel vehicle-peptide molecule or its multimers useful for treating  
 PT inflammatory and autoimmune diseases, cancer, rheumatoid arthritis,  
 PT diabetic retinopathy, obesity, sleep disorders and infertility -  
 PS  
 PS Claim 21: Fig 7; 176pp; English.  
 PS  
 PS The present invention describes a vehicle-peptide molecule (I) or its  
 CC multimers. (I) can have antinflammatory, antitumour, immunosuppressive,  
 CC cytostatic, antitumour, antitachycardic, antidiabetic, ophthalmological,  
 CC antinaemic, anorectic, antinfertility, haemostatic, dermatological and  
 CC neuroprotective activities. (I) can be used as a therapeutic or  
 CC prophylactic agent as well as for screening purposes. (I) is useful for  
 CC diagnosing diseases characterised by dysfunction of their associated  
 CC protein of interest, for identifying normal or abnormal proteins of  
 CC interest, as a part of diagnostic kit to detect the presence of their  
 CC proteins of interest in a biological sample. Additionally, (I) is useful  
 CC for treating inflammatory and autoimmune diseases, tumour growth, cancer,  
 CC rheumatoid arthritis, diabetic retinopathy, obesity, sleep disorders,  
 CC infertility, and neurological degenerative diseases. (I), comprising  
 CC EPO-mimetic compounds are useful for treating disorders characterised by  
 CC low red blood cell levels such as anaemia. The TPO-mimetic comprising  
 CC compounds are useful for treating conditions that involve an existing  
 CC megakaryocyte/platelet deficiency or an expected megakaryocyte/platelet  
 CC deficiency, such as thrombocytopaenia, aplastic anaemia, metastatic  
 CC tumour which result in thrombocytopaenia, systemic lupus erythematosus,  
 CC and Fanconi's syndrome. ABB72403 to ABB73426 and ABL35695 to ABL35777  
 CC represent amino acid and nucleic acid sequences used in the  
 CC exemplification of the present invention.  
 XX  
 XX Sequence 247 AA:  
 SQ  
 Query Match 93.4%; Score 1246; DB 23; Length 247;  
 Best Local Similarity 94.3%; Pred. No. 1.6e-35;  
 Matches 231; Conservative 0; Mismatches 2; Indels 12; Gaps 1;

QY 1 MDKTHCPCPAPPELLGGSVFLFPKPKDTLMISRTPEVTCVVDVSHEDPEVKFNWY 60  
 DB 1 MKDTHCPCPAPPELLGGSVFLFPKPKDTLMISRTPEVTCVVDVSHEDPEVKFNWY 60  
 OY 61 DGEVHNATKPREOYNSTYRVSVLVVLDHMDMLNGKREYKCKVSNKALPAIEKTISKA 120  
 DB 61 DGEVHNATKPREOYNSTYRVSVLVVLDHMDMLNGKREYKCKVSNKALPAIEKTISKA 120  
 OY 121 KGQPREPOVYTLPPSRDELTKNQVSLTCLVKGFYPSDIAVEMESNGQPENNYKTTPPVLD 180  
 DB 121 KGQPREPOVYTLPPSRDELTKNQVSLTCLVKGFYPSDIAVEMESNGQPENNYKTTPPVLD 180  
 OY 181 SGGSFELYSKLTVDKSRMOQGNVSCSVHMEALHNHYTQKSLSLSPGK-----228  
 DB 181 SGGSFELYSKLTVDKSRMOQGNVSCSVHMEALHNHYTQKSLSLSPGK-----228

Db 24278 SSAISVTTCTRIVNLTENAGYYFRVSAQNTFGISDPLEVSSVVIKSPFEKPGAPGKPTIT 24337  
QY 205 -----SCSV----- 208  
|||  
Db 24338 AVTKDSVVWKPASDGGAKIRNYYLEKREKKONKWISVTTETRETFSVKNLIEGLE 24397  
QY 209 ----- 208  
Db 24398 YEFVRKCNLGESEWSEISEPITPKSDVPQAPHFKEELNLRVYQSNATLVCKVTGH 24457  
QY 209 ----- 208  
Db 24458 PRPIVWYRQGEIADGLKYRIOEFKGYHQIILIASVTDDBATVYQVRATNOGGSVST 24517  
QY 209 ----- 208  
Db 24518 ASLEVEVPKAIHLPKTLEGMGAVHALRGEVSIKIPFGKPDVPVTWQKGQDLIDNNGHY 24577  
QY 209 ----- 208  
Db 24578 QVIVTRSFVSLVPPNGVERKDAGFYVYCAKNRFGIDQKTVELDVADVPDPGRGVYSDAS 24637  
QY 209 ----- 208  
Db 24638 RDSVNLWTPEASDGGSKITNIVEKCATTAERWLRVGQARETRYTVINLFCKTSYQFVR 24697  
QY 209 ----- 208  
Db 24698 IAENKFGLSKPSEPSEPTITKEDKTRAMNYDEEVDREVSMTKASHSSTKELYKYMTA 24757  
QY 209 -----MHEA- 212  
|||  
Db 24758 EDLGRGEFVHRCVETSSKKTMAKFKVKGTDQVLVRKEISILNIARHRNILLHESF 24817  
QY 213 -----LHNH----- 216  
|||  
Db 24818 ESMEELVLFETISGLDIFERINTSAFELNREIYSYHVQVCEALQFLSHNIGHDIRP 24877  
QY 217 ----- 216  
Db 24878 ENIIYOTRRSSTIKIIEFGQARQLKPGDNFRLLFTAPEYIYAEVHQHDVVSTATDMSLG 24937  
QY 217 -----YT----- 218  
|||  
Db 24938 TLVVLLSGINPLAETNOQIIENINNAEYTFDEAFKIEIAMDVDRLLVKKRKSAM 24997  
QY 219 -----OKS 221  
|||  
Db 24998 TASEALQHPWLKQIERSVSTKIVIRTLKHRRYYHTLLIKDLNMVVSARISCGGATRSQKG 25057  
QY 222 LSL----- 225  
|||  
Db 25058 VSAKVVKVASIEIGPVSGOIMHABEGGHVKYVCKIENYDQSTQVWTWFGVRQLENSEK 25117  
QY 226 ----- 225  
Db 25118 YEITYEDGVAILLYVKDITKLDDGTVRCKVYNDYGEDSSYAEFLVKGREVDYCYRRTMK 25177  
QY 226 ----- 225  
Db 25178 KIKRRTDMRLERPEFTLPLYNKTAIYGVENVRGVTTIVHPEPHVTWYKSGQIKPGD 25237  
QY 226 ----- 225  
Db 25238 NDKKTYFESDKGLYQLTINSVTTDDAEYTVVARNKYGEDSCKAKLTVLHPPTDSTLR 25297  
QY 226 ----- 225  
Db 25298 PMFKRLLANEQEGSGVCFEIRVSGIPPTLKWEDGQPLSIGPNIEIIEHGLDYALH 25357  
QY 226 ----- 225  
Db 25358 IRDTLPEDTGYRVTAINTAGSTSCQAHLOVRLRYKKOEFKSKEEHHRVQKIDKTULR 25417

QY 226 ----- 225  
Db 25418 MAEILSGTESVPLTQVAKREALREAAVLYKPAVSTKTVKGEFRLEIEEKKERKLMPYDV 25477  
QY 226 -----PGK----- 228  
|||  
Db 25478 PEPRKYKOTTIEEDORIKOFVPMDSMKWKIKINDQYEMFGKLDVRVQVKRPKIRLSWBQ 25537  
QY 229 ----- 228  
Db 25538 FYVMPLPRTDOYRKPWRIPKLSODDLEIVRPARRTPSPDYDFYRPRRSISLGDISDEE 25597  
QY 229 ----- 228  
Db 25598 LLLPIDDYLAMKRTBEERLURLEEELELGFSPSPSPHFEELSRLYSYSPQAHVKEET 25657  
QY 229 ----- 228  
Db 25658 RKNFRYSTVHIPTKAEASTSYAELRERHAQAAAYRQPKQRIMAREDEBELLRPVTTQH 25717  
QY 229 ----- 228  
Db 25718 LSEYKSELDPMSEKSRKSRQRREVTEITEEBEYFISKHAQRESSSASRLLRRRS 25777  
QY 229 ----- 228  
Db 25778 LSPVTIELMRPVSELIRSRQPAEEYEDDTERRPPTPRSPSPVSSERSLSRFRS 25837  
QY 229 ----- 228  
Db 25838 ARFDIFSRYESMKAALKTKQTSERKEYVLSQQPFTLDHAPRITLRMRSHRVPQGNTRFI 25897  
QY 229 ----- 228  
Db 25898 LNVQSKPTAEVKWYHNGVELQESSKIHYTNTSGVLTLEILDCHTDDSGTYRAVCTNYKE 25957  
QY 229 ----- 228  
Db 25958 ASDYATLDVTGGDYTYASQRRDEEVPSPFPELRTAEYAVSPFKKTSMEASSSVREV 26017  
QY 229 ----- 228  
Db 26018 KSQMTETRESLSYEHSAEMKSAALEKESLEKSTTRKIKTTLAARILTKPRSMTVYE 26077  
QY 229 -----DWL----- 231  
|||  
Db 26078 GESARFSCDTPGEPVPTVTLRKGOVLSTSAHQVTTTKYSTFEISSVQASDEGNYSV 26137  
QY 232 ----- 231  
Db 26138 VENSEGQBAEFTLTIQKARVTEKAVTSPPRVKSPPEPRVKSPKAVKSPKRVKSPSPHPK 26197  
QY 232 ----- 231  
Db 26198 AVSPTETKTPREKVOHLPVSAPPKITQFLKABASKEIAKLTCVWESSVLRAKEVTWYKD 26257  
QY 232 -----KAF 234  
|||  
Db 26258 GKLLKENGHFQHYSDAGTYELKINNLTESDQGEYVCEISGEGGTSKTNLQPMGOAFKSI 26317  
QY 235 YDKVAE----- 240  
|||  
Db 26318 HEKVSКИSETKKSDQKTTESTVTTRTEPKAPEPISSKPVITGLQDQTTVSSDSVAKFAVK 26377  
QY 241 ----- 240  
Db 26378 ATGEPRPTAIWTKDGAITQGGKYKLSDEKGGFFLEIHKTDTSDSGLYCTYTKVNSAGSVS 26437  
QY 241 -----KLKEA 245  
|||  
Db 26438 SSCKLTIKAIKTEAQKVSTQKTSETITPOKAVVQBEISQKALRSEIKMSEA 26490

Qy 183 ----- 182  
Db 22118 EFRVAKNDAGYSEPREAFSSVIIKEQIETADLTGITNLIITCKAGSPFTIDVPISGR 22177  
Qy 183 -----GSFPL----- 187  
Db 22178 PAPKVTWLEEMRLUKETDRVSITTTKORTTLTVKDSMRGDSGRYFLTLENTAGVKTFEVS 22237  
Qy 188 ----- 187  
Db 22238 VVIGRPGVPTGPIEVSSVSAESCVLSWGEPKDGGEITNVIYVEKRESGTTAMQOLVNS 22297  
Qy 188 ----- 187  
Db 22298 VKRTQIKVHTLTMYEYFRVSSENREGVSKPLESAPIIAHPPFPAPSAPTRPEVYHSA 22357  
Qy 188 ----- 187  
Db 22358 NAMSIRWEEPYHDGSKIIGYWEKKEKERTILWVKENKVPCECNKYVTGLVEGLEQFR 22417  
Qy 188 ----- 187  
Db 22418 TYALNAGVSKASEASRPIMAQNPVADAPRPEVTDTRSTVSLIWSAPAYDGGSKVVGVI 22477  
Qy 188 ----- 187  
Db 22478 IERKPVSEVGDGRLKCNVTIVSDNFFTVTALSEGDYEFVRVLAKNAGVISKSESTGP 22537  
Qy 188 ----- 187  
Db 22538 VTCREYAPPAELDARLHGLVTRAGSDLVDAAVGGKPEPKIITWTKGDKELDLCEKV 22597  
Qy 188 ----- 187  
Db 22598 SLOYTKRATAVIXPCDRSDSGKYTLTVKNASGTVKAVSMVKVLDSPGCKLTVSRVQ 22657  
Qy 188 ----- 187  
Db 22658 EKCTLANSLPOEDGGAETHYIVERRETSRLNWIVIEGECTLSYVVTRLIKNEYIFRV 22717  
Qy 188 ----- 187  
Db 22718 RAVNKYGPVPESEPIVARNSTIPSPGPIPEVGTGKEHIIQWTKPESDGGNEISNY 22777  
Qy 188 ----- 187  
Db 22778 LVDRREKESLRWTRVNKDYYVYDRLKVTSLMEGCDYQFRVTAVNAAGNSEPSPERSNFIS 22837  
Qy 188 ----- 187  
Db 22838 CREPSYTPGPPSAPRVVDTTKHSISLAWTKPMYDGGTDIVGYVLEMQEKOTDQWYRVHTN 22897  
Qy 188 ----- 187  
Db 22898 ATIRTEFTVPLKMGOKYFRVAANVKGMYSESIETEPVERIEIPDELEADOLKK 22957  
Qy 188 -----YSKLTVDK----- 195  
Db 22958 TVTTRAGASLRLMVSGRPPPVITWSKQIDLASRAIIDTTESYSLLIYDKVNRDAGK 23017  
Qy 196 ----- 195  
Db 23018 YTIEAENQSGKSATVLVKYVDTPGPCPSVKVKEVSRDSVTITWIEPTIDGAPINNIY 23077  
Qy 196 ----- 195  
Db 23078 EKREAAAMRAKFTVTKCKSKTLYRISGLVEGTMHYFRVLPENIYIGBPCETSDAVLVSEV 23137  
Qy 196 ----- 195  
Db 23138 PLVPAKLEVVDVTKSTVTLAWEKPLXDGGSLRTGYVLEACKAGTERMKVVTLKPTVLEH 23197  
Qy 196 ----- 195

Db 23198 TVTSLNEGEQYLFRIQANEKGVSEPRETVTAVTVQDLRVLPTIDLSTMPQKTIHVDPAGR 23257  
Qy 196 ----- 195  
Db 23258 PVELVIPAGRPAPPAASWFFAGSKLRESERVVETHTKVAKLTIRETTRDTGYEYTLLEK 23317  
Qy 196 ----- 195  
Db 23318 NVTGTTSETIKVIIIDKPGPPTGPIKIDEIDATISITISWEPPELGGAPLSGYVVEORDA 23377  
Qy 196 ----- 195  
Db 23378 HRCGLPVSBSVTRSTFKFTRLTEGNEVYFRVAATNREGISYVLSQSEVIECRSIRIPGP 23437  
Qy 196 ----- 195  
Db 23438 PETLIQIFDVSRDGMLTWYPPEDDQSGQVTGYIVERKEVRADRWVRVKNKVPVMTYRST 23497  
Qy 196 ----- 195  
Db 23498 GLTEGLEYHRVTRAINARGSGKPSRPSKPIVAMDPIAPPKQPONPRVTDTRTTSVSLAWS 23557  
Qy 196 ----- 195  
Db 23558 VPEDEGSKVTGYLIEQKVDOHEWTKCNTPTTKIREVTLTLHPQGAERYRVLACNAGG 23617  
Qy 196 -----SRWQOG----- 201  
Db 23618 PGPAEYVPGTVKVTMELEYPDYELDERYQEGIFVRQGGVIRLTIPIKGPFPICKWTKEG 23677  
Qy 202 ----- 201  
Db 23678 QDISKRAMIATSETHTELVIKEADRGDSGYDVLNKKCGKAVYIKVRIGSPNSPEGP 23737  
Qy 202 ----- 201  
Db 23738 LEYDDIOQVRSVRVSWRPPADDDGADILGYILEREVEPKAAWYIDSVRGTSLVVKGLE 23797  
Qy 202 ----- 201  
Db 23798 NVEYHFRVSAENOFGISKPLKSEPVTPKPLNPPEPPSNPPEVLDVTKSSVLSWSRPK 23857  
Qy 202 ----- 201  
Db 23858 DGGSRVTGYIYIERKETSTOKVVRHNKTQITTTWYTVGLVPDAEYQFRIIAQNVDGLSE 23917  
Qy 202 ----- 201  
Db 23918 TSPASEPVCKDPDKPSQPGELEILSISKDSVTLOWEKPECDDGKELGWWVEYRQSGD 23977  
Qy 202 ----- 201  
Db 23978 SAWKSNKERIKDKQFTIGGLEATEYFRVFAENETGLSRPRRTAMSIKTLTSGEAPG 24037  
Qy 202 ----- 201  
Db 24038 IRKEMKDVTTKLGEEAOLSCQIVGRPLPDIKWYRFGKELIOSRKYKMSDGRHTLTVMT 24097  
Qy 202 ----- 201  
Db 24098 EEQDEGEVYTCIATNEVEGETSSKLLQLQATPQFHPGCPYPLKEYYGAVGSTLRHLVWYIG 24157  
Qy 202 -----NVF----- 204  
Db 24158 RVPAMTWFHQKLLQNSENITIENTBHYTHLVKMNQVRKTHAGKYKQLSNVEGTDAI 24217  
Qy 205 ----- 204  
Db 24218 LDVEIODKPKPTGPIVIEALLKNASVSWKPPADGGSGITNVVVEKCEAKEAEQOLV 24277  
Qy 205 ----- 204

Db 19898 WLKPEHGGSRIVHYVVEALEKGQKNWKCVAKSTHHVVSGLRENSEYFFRVFAENOAG 19957  
QY 176 ----- 175  
Db 19958 LSDPRELLLPVLKKEOLEPPEIDMKNFSPSHYVVRAGSNLKYDIPISGKPLPKVTLSRDG 20017  
QY 176 ----- 175  
Db 20018 VPLKATMRENTETAENLTINLKESVTADAGRYEITAANSOGTTKAFINIVVLDPRGPPT 20077  
QY 176 ----- 175  
Db 20078 GPWISDITEESVTLKNEPPKYDGGSQVTNYILLKRETSTAVWTEVSATVARTMKVMKL 20137  
QY 176 ----- 175  
Db 20138 TTGEYQFRIKAENRFGISDHIDSACVTVKLPYTPPGPSTPWVTVNVTRESITVGHEPV 20197  
QY 176 ----- 175  
Db 20198 SNGSAVGYHLEMKDRNSILWQKANKLVIRTHFKVTTISAGLIYEFVRYAENAGVGK 20257  
QY 176 ----- 175  
Db 20258 PSHPEPVLADACEPPRNRVITDISKNSVLSWOQAPADGGSKITGYIVERDDLDPGRW 20317  
QY 176 ----- 175  
Db 20318 TRASFNTVTEQTTISGLTONSQYEFVRFARNAGVSNPSEVVGPIITCIDSYGGVIDL 20377  
QY 176 ----- 175  
Db 20378 PLEYEVVKYRAGTSVKLRAGISGKPAPTIEWYKDDKELQTNALVCVENTTDLASILKD 20437  
QY 176 ----- 175  
Db 20438 ADRLNSGCYELKLRNMAASATIRVQILDKPGPGPIEFKVTVAEKTILLWRPPADDG 20497  
QY 176 ----- 175  
Db 20498 GAKITHYIVEKRETSRVVMSVSEHLEECIITTKIIRGNEYIFRVAVNKYIGIGEPLES 20557  
QY 176 ----- 175  
Db 20558 DSVAKNAFVTPGPGIPEVTKITNSMTVWSRPIADGGDISGYFLEKRDKKSIGWFK 20617  
QY 176 ----- 175  
Db 20618 VLKETIRDTRQKVTGLTENSQYQVCAVNAAGQGFSEPFSEFYKAADPIDPPGPAKIR 20677  
QY 176 ----- 175  
Db 20678 IADSTKSSITLGWSKPVYDGGSAVTGYVVEIROGEEETVSTRKEVTRTEYVVSNLKP 20737  
QY 176 ----- 175  
Db 20738 GYNYFRVSAVNCAGGEPPIEMNEPVQAKDILEAPEIDLVALRTSVIAKAGEDVQVLIP 20797  
QY 176 ----- 175  
Db 20798 FKGRPPPTVTRKDEKLGSDARYSIENTDSSLLTIPQVTRNDTGKYLITTIENGVEPK 20857  
QY 176 ----- 175  
Db 20858 SSTSVKVLDTPAACQKLOVKHVSRTVTLWDPLIDGSPINIVYIEKRDATKRTWSV 20917  
QY 176 ----- 175  
Db 20918 VSHKCSSTFKLIDLSEKTPFFFRVLAENEIGIGPCETTEPVKAAEVPAPTRDLSMKDS 20977  
QY 176 ----- 175  
Db 20978 TKTSVILSWTKPFDGGSVITEYVVVERKKGQBTWSHAGISKTCIEIYVSOLKEQSVLEFR 21037

QY 176 ----- 175  
Db 21038 VFAKNEKGLSDPVTGPIITVKELIITPEVDLSDIPGAQVTVRIGHNVHLELPYKGPKEPS 21097  
QY 176 ----- 175  
Db 21098 ISWLKDGFLPKESEFVRFSTENKITLSIKNAKHEGGKYTVILDNAVCRIAVPITVITL 21157  
QY 176 ----- 175  
Db 21158 GPSPKPKPIREFDEIKADSVILSWDPEDNGGGEITCYISIEKRETSQTNWKMVCSSVART 21217  
QY 176 ----- 175  
Db 21218 TEKVPNLKDAEYQFVRAENRYGVSQPLVSSIIVAKHQFRIPPGPKGVINYNTSDGMS 21277  
QY 176 ----- 175  
Db 21278 LTWDAVPYDGGSEVTGFHVEKKERNISILWQKNTSPISGREYRATGLVBGLDYQFRVYAE 21337  
QY 176 ----- 175  
Db 21338 NSAGLSSPDPKSKFTLAVSPVDPPTDYIDVTRETITLKNPPLRDGSGKIVGYSIEXR 21397  
QY 176 ----- 175  
Db 21398 QGNERWRCNFTDVSECOYTVTGLSPGDRYEFRIIARNAVGTISPSPSSQSGIIMTRDENV 21457  
QY 176 PPV ----- 178  
Db 21458 PPIVEFGPEYFDGLIJKGESLRIKALVOGRVPVPRVTFKOGVEIEKRMNMEITNVLGST 21517  
QY 179 ----- 178  
Db 21518 SLFVRDATPDHRGVTVVEAKNAGSAAELKVKVQDTPGKVVGPIRFTNITGEKMTLWWD 21577  
QY 179 ----- 178  
Db 21578 APLNDGCAPITHYIIEKRETSRLAWALIEDKCEAQSYTAIKLINGNEYQFRVSAVNKFGV 21637  
QY 179 ---LDS----- 182  
Db 21638 GRPLSDPVAQIQVTPDAPGPIEPSNITGNSIILTWARPESDGGSEIQOILEREKK 21697  
QY 183 ----- 182  
Db 21698 STRWVKISKRPISETRFPKVTGLTEGNEYEFHVMAENAAAGVGPASGISRLIKCREPVNPP 21757  
QY 183 ----- 182  
Db 21758 GPPTVVKVTDTSKTTVSLWSKPVFDGGMELIIGYIEMCKTDLGDWHKVNABACVKTRYT 21817  
QY 183 ----- 182  
Db 21818 VTDLOAGEEYKFRVSAINGAGKGDSCVETGTIKAVDRLTAPELDIDANFKQTHVVRAGAS 21877  
QY 183 ----- 182  
Db 21878 IRLFIAYQGRPTPTAVWSKPDNSLSLRADIHTTDSFSTLTVCNCRNDAGKYTLTVENNS 21937  
QY 183 ----- 182  
Db 21938 GSKSITFTVKVLDTGPPGPIITFKDVTGRSATLMDAPLLDGGARIHHYVVEKREASRS 21997  
QY 183 ----- 182  
Db 21998 WQVISEKCTRQIFKVNDLAEGVPYFVRSVAVNEYGVGEPEYEMPEPIVATEQAPPRRLDV 22057  
QY 183 ----- 182  
Db 22058 VDTSSSAVLAWLKPBDHGGSRITGYLLEMRQKGSGLWVEAGHTKOLTFTVRLVEKTEY 22117

Qy	171	-----	170
Db	17738	KPEHGGSRVLGYVVEMQPKTEKWSIVAESKVCNAVVTGLSSGOEYQFRVKAYNEKGKS	17797
Qy	171	-----	170
Db	17798	DPRLGVPIAKDLTIQPSLKLPTNTYSIQAGEDLKIEIPVIGRPRPNISWVKDEPLKQ	17857
Qy	171	-----	170
Db	17858	TTRVNETATSTVLHIKEGNKDDGKYTVTATNSAGTATENLSVIVLEKPGVPVGRVF	17917
Qy	171	-----	170
Db	17918	DEVSADFVVISWEPPAYTGGCQISNYIVEKRDTTTTTWHMVSATVARTTIKITLKTGTE	17977
Qy	171	-----	170
Db	17978	YQFRIFAENRYKSAPLDSKAVIYOYPEKPGPTPEVTSISKQMDLVOMHEPVNDGGT	18037
Qy	171	-----	170
Db	18038	KIIGYHLEQKEKNSTLWVKLNKTIQDTKFKTTGLDCELEVEFKVSAENIVGIGKPSKVS	18097
Qy	171	-----	170
Db	18098	ECFVARPCDPPRPEAIVITRNNVTLWKKKPAYDGGSKITGYIVEKKDLPGRWMKASF	18157
Qy	171	-----	170
Db	18158	TNVLETEFTVSGLEDQRYEFVRVIAAAGNFSEPSDSSGAITARDEIDAPNASLDPKYK	18217
Qy	171	-----	170
Db	18218	DVIVVHAGETVLEADIRGKPIPDVWMSKCKELEETAARMEIKSTIOKTTLVVKDCIRT	18277
Qy	171	-----	170
Db	18278	DGGYIILKSNVGGTKSIPITVKVLDRPGSPGLKVTGVTAEKCYLAWNPLODGGANI	18337
Qy	171	-----NKYTT-----	175
Db	18338	SHYIIEKRETRSLSWTQVSTEQVALNKYTKLLPGNEYIFRVMVANKYGIGEPLESGPVT	18397
Qy	176	-----	175
Db	18398	ACNPYKPGPSTPEVSATIKDSMVVTWARVPDGGTEIEGYILLEKRDKEGVRWTKCNKK	18457
Qy	176	-----	175
Db	18458	TLTDLRLRVTLGTGHSYEFRAAENAAGVGEPSVIFYRACDALYPPGPPSNPKVTDIT	18517
Qy	176	-----	175
Db	18518	SRSSVSLAWSKPIYDGCAPVKGYVVEVKEAADEWTTCTPTGLQGKQFTVTKLKENTFY	18577
Qy	176	-----	175
Db	18578	NFRICAINSEGVGPATILPGSVQAQRIEPPIEILDADLRKVVVLRASATLURLFTWTKGR	18637
Qy	176	-----	175
Db	18638	PEPEVKWEKABGILTDRAQIEVTSFTMLVIDNTRFDGRYNLTLNENSGSKTAFVNVR	18697
Qy	176	-----	175
Db	18698	VLDSPSAPVNLTIREVKKDSVTLSEWPEPLIDGAKITNYIVEKRETRKAYATITNNCTK	18757
Qy	176	-----	175
Db	18758	TTFRINLQEGCSYFRVLASNEYIGLPAETTEPVKVSPEPLPPGRVTLVDVTRNTATI	18817
Qy	176	-----	175

D	b	18818	KWEKPE	DGSGKITGVVYEMOTKGSEKSWSTCTQVKTLBATISGLTAGEEYVFRVAANNEK	18877
Q	y	176	- - - - -	- - - - -	175
D	b	18878	GSRDPRQLGVPVIARDIEIKFSPVELPPHTFNVKAREQLKDIDVPEKRPQATVNWRKDGQT	18937	
Q	y	176	- - - - -	- - - - -	175
D	b	18938	LKETTRNVSSKTVTLSIKEASKEDYGTYELCVSNSAGSITPITIIVILDRPGPPGPPI	18997	
Q	y	176	- - - - -	- - - - -	175
D	b	18998	RIDEVSCDSIITISWNPDEYGGCOISNIVEKKETTSTTWHIVSOAVARTSIKIVRLTTG	19057	
Q	y	176	- - - - -	- - - - -	175
D	b	19058	SEYQFRVCAENRYGSSYESAVVAEYPFPSPGPGTPKVHHATKSTMLTMTQVPVNDG	19117	
Q	y	176	- - - - -	- - - - -	175
D	b	19118	GSRTGYHLEYKERSIILWSKANKILLIADTOVKVSGLDEGLMEYRVYAENIAGIGCKSK	19177	
Q	y	176	- - - - -	- - - - -	175
D	b	19178	SCEPVPARPCDDPCQPPEVTNITRKSLSLKWKSPHYDGGAKITGYIVERRELDPGRWLKC	19237	
Q	y	176	- - - - -	- - - - -	175
D	b	19238	NYTNIOETYFEVTELTEDORYEFVRFAANAADVSEPSESTGPIIVKDDVEPPPVMMDVK	19297	
Q	y	176	- - - - -	- - - - -	175
D	b	19298	PRDIVVKAGEVLKINADIARPLPVISWAKDGTIEERARTEIISTDNHTLLTVKDCIR	19357	
Q	y	176	- - - - -	- - - - -	175
D	b	19358	RDTOGYVLTLNKNAVGRSVAVNCVKLDKPGPPAGPLEINGLTAEKCSLSWGRPQBDGGAD	19417	
Q	y	176	- - - - -	- - - - -	175
D	b	19418	IDYYHRKKRETSHLAWTICEGELOWTSCKVTLLKGNEYIFRVTGVNKYGVGEPLSVAI	19477	
Q	y	176	- - - - -	- - - - -	175
D	b	19478	KALDPFTVPSPPTSLIETSVTKESMTLCWRSPESDGGSEISGYIIERREKNLSRWVRYNK	19537	
Q	y	176	- - - - -	- - - - -	175
D	b	19538	KPYVDLRVKSTGLREGCEYEYRVYAENAAGLSLPSETSPLIARAEDPVPSPSPKPPIVD	19597	
Q	y	176	- - - - -	- - - - -	175
D	b	19598	SGKTTITIAWVKPLPDGCAPITGYTVYKKSDDTDKWTSIQSLRGTEYTIISGLTTGAEYV	19657	
Q	y	176	- - - - -	- - - - -	175
D	b	19658	FrvksvNkVGASDPSSSDPOIAKEREBEPLFDIDSEMKRKLIVKAGASTMTVPFFRGRP	19717	
Q	y	176	- - - - -	- - - - -	175
D	b	19718	VPNVLSKPDPTDLRTRAYVDITDSRTLST IENANRNDSGKYTLTIONVLSAASLTLVVKV	19777	
Q	y	176	- - - - -	- - - - -	175
D	b	19778	LDTPGPTNITVQDYTKESA VLSDVPENDGGAPVKYNHIEKREASKAWYSVTNNC NRL	19837	
Q	y	176	- - - - -	- - - - -	175
D	b	19838	SYKVTNLQEGAIYFRVRSGENEFGVIPAETKEGVKITEKPSPEKLGVTISIKDSVS LT	19897	
Q	y	176	- - - - -	- - - - -	175

Db 15518 LOEGCSYFRVLAENEYIGLPAETAESVKASERDLPPCKITLMDVTRNSVLSWEKPEH 15577  
QY 168 ----- 167  
Db 15578 DGGSRILGYIVEMQTKGSDKWATCATVKTEATITGLIQEEYSFRVSAQNEKGISDPRQ 15637  
QY 168 ----- 167  
Db 15638 LSPVIAKDLVIPPAPKLLFNFTTFLAGEDLKVDVFFIGRPTPAVTWHKDNVPLKQTTVR 15697  
QY 168 ----- 167  
Db 15698 NAESTENNSLLTIKDACREDVGHYVVKLTNSAGEAIELNVILDKPGPPTGPMDEVT 15757  
QY 168 ----- 167  
Db 15758 ADSITLSWGPYPKYDGGSSINNIYIVEKRDITSTTTWQIVSATVARTTIKACRLKTGCEYQFR 15817  
QY 168 ----- 167  
Db 15818 IAAENRYGKSTYLNSEPTVAQYFPKVPGPPTVVTLSSRDSMEVQWNEPISDGGSRVIG 15877  
QY 168 ----- 167  
Db 15878 YHLERKERNILWKLKNTPIPTQTKFTGTLEEGVEYEFVSAENIVIGIKPSKVSECV 15937  
QY 168 ----- 167  
Db 15938 ARDPCDPPGRPEAIIIVTRNSVTLOWKKPTIYDGGSKITGYIVEKKELPEGRWMAKSFNTII 15997  
QY 168 ----- 167  
Db 15998 DTHFEVTGLVEDHRYEFVRIARNAAGVFSEPSSTGAITARDEVDPPIRISMDPKYKDTIV 16057  
QY 168 ----- 167  
Db 16058 VHAGESFKVDADIYKPIPTIOWIKGDQELSNTARLEIKSTDFATSLSVKDAVRVDSGN 16117  
QY 168 ----- 167  
Db 16118 ILKAKNVAGERSVTYVNVKLDPRPGPVVISGYTAEKCTLAWKPPLODGGSDIINIV 16177  
QY 168 ----- 167  
Db 16178 ERRETSRLVWTVVDANVOTLSCKVTKLLEGNEYTFRIMAVNKYGVGEPLESPVAKNPF 16237  
QY 168 ----- 167  
Db 16238 VVPDAPKAPETVTYTKDSMIVVWERPASDGGSEILGYVLEKRDKEGIRWTRCHKRLIGEL 16297  
QY 168 ----- 167  
Db 16298 RLVRTGLIENHDEYFRVSAENAGLSEPPSAYQKACDPIYKPGPPNPKVIDITRSSV 16357  
QY 168 ----- 167  
Db 16358 FLWSKPIYDGCETIQGYIVEKCDVNVGEWMTCTPTTGINKNTIEVEKLEKHEYNFRIC 16417  
QY 168 ----- 167  
Db 16418 AINKAGVGEHADVPGLIIEEKEAPDIDLDELKRIINIRAGGSLRLFPVPIKGRPTPEV 16477  
QY 168 ----- 167  
Db 16478 KMGKVGDEIRDAAIIDVTSSTSLVDNVNRYDSCKYTLTLENSSGTKSAFVTVRVRLDTP 16537  
QY 168 ----- 167  
Db 16538 SPPVNLKTEITKDSVITWEPLLDGGSKIKNYIVEKREATRKSAAVVTNCHKNWKI 16597  
QY 168 ----- 167  
Db 16598 DQLQEGCSYFRVTAENEYIGLPAQTADPIKVAEVPQPGKITVDDVTRNSVLSLWTKP 16657

QY 168 ----- 167  
Db 16658 EHDGSKIIQYIVEMQAKHSEKSEKARVKSILQAVITNLTOGEEYLFVRVAVNEKGRSDP 16717  
QY 168 ----- 167  
Db 16718 RSLAVPIVAKDLVIEPDVKPAFSSYSVOVGDLKMEVPISGRPKPTITWTKGLPLKQTT 16777  
QY 168 ----- 167  
Db 16778 RINVTDSLDLTLTSIKETHKDDGGGYITVANVWCQTASIEIVTLDRKDDPKPGVKFDD 16837  
QY 168 ----- 167  
Db 16838 VSAESITLSWNPPLYTGCOITNIVQKRDTTTTVMVDVVSATVARTTLKVTCLKTGTETVQ 16897  
QY 168 ----- 167  
Db 16898 FRIFAENRYGOSFALESDFIVAQYPKPEPGTPTAFATAISKDSMVIQWHEPVNNGGSPV 16957  
QY 168 ----- 167  
Db 16958 IGYHLERKERNILWKLKNTIIHDTQFKAQNLBEGIEYEFVRYAENIVGVGKASKNSEC 17017  
QY 168 ----- 167  
Db 17018 YVARDPCDPPGTPEPIWVKRNEITLQWTKPVYDGGSMITGYIVEKRDLPDGRWMAKSFN 17077  
QY 168 ----- 167  
Db 17078 VIETQTVSGLTEDQRYEFVRIAKNAAGAIKSPSDSTGPTAKDEVLPRIKMDPKFRDT 17137  
QY 168 ----- 167  
Db 17138 IVNAGETFRLEADVHGRPLPTIEWLRGDKEIEESARCEIKNTDFKALLIVKDAIRDGG 17197  
QY 168 ----- 167  
Db 17198 QYILRASNVAGSKSPVNVKVLDRPGPEPVQVTGVTSEKSLTWSPPLQDGGSDISHY 17257  
QY 168 ----- 167  
Db 17258 VVEKRETSRLAWTVVASEVVTNSLKVTKLLEGNEYVFRIMAVNKYGVGEPLESAPVLMKN 17317  
QY 168 ----- 167  
Db 17318 PFVLPCKPSLEVTNIAKDSMTVCWNRDPDSDGSGSEIIGYIVEKRDGRGIRWIKCNKRIT 17377  
QY 168 ----- 167  
Db 17378 DLRLRVTLGTEDHEYEFRVSAENAGVGEPSPATVYKACDPVFKPGPPTNAHIVDTTKN 17437  
QY 168 ----- 167  
Db 17438 SITLAWKPIYDGSSEILGYVVEICKADEEEMQIVTPQTLRVTRFEISKTEHOEYKIR 17497  
QY 168 ----- 167  
Db 17498 VCALNKVGLGEATSVPGTVKPEDKLEAPELDLDELKRIINIRAGGSLRLFPVPIKGRPMP 17557  
QY 168 ----- 167  
Db 17558 ETTWSREGEFTKVQIEKGVNYTQLSIDNCNDAGKYILKLENSGSKSAFVTVKVLVD 17617  
QY 168 ----- 170  
| : |  
Db 17618 TPGPPONLAVKEVRKDSAFLVWEPPIIDGGAKVKVNYIDKRESTRKAYANVSSKCSKTSF 17677  
QY 171 ----- 170  
Db 17678 KVENLTEGAIYFRVMAENFVGVPVETVDVAKAEPSPKVTLTLDVTSQTSASLMWE 17737

QY 161 ----- 160  
Db 13358 SYFRVFAENYEGIDPGETRDVAKASOTPGPVVDLKVRSVSKSSCSIGHKKPHSDGSR 13417  
QY 161 ----- 160  
Db 13418 IIGYVVDFLTEENKQRMKSLQISAKDLTEGKEYTFRVSAENEGECTPSEITVVAR 13477  
QY 161 ----- 160  
Db 13478 DDVVAPDLUKLPDLCLYLAKENSFRKLPIKGPAPSVSWKKGEDPLATDTRSVSS 13537  
QY 161 ----- 160  
Db 13538 AVNTTLIVDCQSDACKYITILKNVAGTKEGTISIKVVKPGIPTGPIKDFDEVTAEAMT 13597  
QY 161 ----- 160  
Db 13598 LKWAPPKDDGGSEITNYILEKRDVANKWVTCASAVQKTFRTVRLHEGMEYTFRVSAN 13657  
QY 161 ----- 160  
Db 13658 KYGVGELKSEPIVARHPFVDPDAPPPNIVVRHDSVSLTWTDPKKTGGSPITGYHLEF 13717  
QY 161 ----- 160  
Db 13718 KERNSLLWKRAKTPIRMRFKVTGLTEGLEYEFRVMAINLAGVGKPSLPSEPVVALDPI 13777  
QY 161 ----- 160  
Db 13778 DPPGKPEVINITRNSVTLIWTEPKYDGHKLTGYIVEKRDLPKSNWKAHVNVPECAFT 13837  
QY 161 ----- 160  
Db 13838 VTDLVEGGKYEFRIRAKNTAGAISAPSESTETIICDEYEAPTIVLDPTIKDGLTIKAGD 13897  
QY 161 ----- 160  
Db 13898 TIVLNAISILGKPLPKSSWSKAGDIRPSDITQITSTPTSSMLTIKYATRKDAGEYITIA 13957  
QY 161 ----- 160  
Db 13958 TNPFGTKVEHVKTIVLDVPGPGPVEISNVSAEKATLTWTPLEDGGSPIKSYILEKRET 14017  
QY 161 ----- 160  
Db 14018 SRLLTWVSEDIOCRHVATKLIQNEIYFRVSAVNHYGKEPVQSEPVKWDVRFPGPGP 14077  
QY 161 ----- 160  
Db 14078 PEKPEVSNVTKNATVSWKRPVDDGGSEITGYHVERREKSLWVRAIKTPVSDLRCKVT 14137  
QY 161 ----- 160  
Db 14138 GLOEGSTYFRVSAENRAGIGPPEASDSVLMKDAAYPPGPPSNPHVTDTRKKSASLAWG 14197  
QY 161 ----- 160  
Db 14198 KPHYDGGLEITGYVVEHQVGDEAWIKDTGTALRITQFVVDLQTKKEYNFRISAINDA 14257  
QY 161 ----- 160  
Db 14258 GVGEPAVIPDVEIVEREMAPDFELDAELRRTLVRAGLSIRIFVPIKGRPAPEVTWKDN 14317  
QY 161 ----- 160  
Db 14318 INLKNRANIENTESFTLLIPECNRYDTGKFVMTIENPAGKSGFVNVRLDTPGVLNL 14377  
QY 161 ----- 160  
Db 14378 RPTDITKDSVTLHWDLPLIDGGSRTINYIVEKREATRKSYSSTATTKCHKCTYKVTGLSEG 14437  
QY 161 ----- 160

Db 14438 CEYFFRVMARENXYGIGEPTTETTEPVPKASEAPSPDLSLINDITKSTVSLAWPKPHDGG 14497  
QY 161 ----- 160  
Db 14498 KITGYVIEAQKSGDQWTHITTVKGLECVVRNLTEGEYTFQVMVNSAGRSAPRESRPV 14557  
QY 161 ----- 160  
Db 14558 IVKEQTMLPELDLURGIYQKLVIAKAGDNKVEIPVGRPKPTVTKKGGDQILKQTORVNF 14617  
QY 161 ----- 160  
Db 14618 ETTATSTILNINECVRSDSGPPLTARNIVGEVDVITIQVHDI PGPTGPIKDFEVSSD 14677  
QY 161 ----- 160  
Db 14678 FVTFSWDPDPPDGGVPISNVVEMROTSTWVELATTVIRTYYKATRLTTGLEYOFRVK 14737  
QY 161 ----- 160  
Db 14738 AQNRYGVGPGITSAMIVANYPFKVPGPPGTPOVTAVTKDSMTISWHEPLSDGGSPILGYH 14797  
QY 161 ----- 160  
Db 14798 VERKERNGILQVTSKALVPGNIFKSSGLTDGTAIEFRVIAENWACKSPKSPSEPMAL 14857  
QY 161 ----- 160  
Db 14858 DPIDPPGKPVPLNITRHTVTLKWAPEYTGCGFKITSYIVEKRDLPNGRMLKANFSNILEN 14917  
QY 161 ----- 160  
Db 14918 EFTVSGLTEDAAYEFRIAANAAGAI SPSPSDAITCRDDVEAPKIKVDVKFKDVILK 14977  
QY 161 -----EWESNGO----- 167  
Db 14978 AGEAFRLADVSGRPPTMEWSKDGKELEGTAKLEIKIADFTNLVKNKDSRRDSGAYTL 15037  
QY 168 ----- 167  
Db 15038 TATNPGFAKHIFNVKVLDRPGPEPLAVTEVTSEKCVLSWFPPLDDGGAKIDHIVYQK 15097  
QY 168 ----- 167  
Db 15098 RETSRLAWTNVASEVQVTKLVTKLLKGNEYIFRVMVANKYGVGEPESEPVLA VNPYGP 15157  
QY 168 ----- 167  
Db 15158 POPPKNPEVTITTKDSMVVCMWGHPSDGGSEIINYIVERRRKAGORWIKCNKKTLDLRY 15217  
QY 168 ----- 167  
Db 15218 KVSGLTEGHEYEFRI MAENAGISAPSPFPYKACDVFVKPGPGPNRVLDTSRSSISI 15277  
QY 168 ----- 167  
Db 15278 ANKKPIYDGGSEITGYVWEIALPEDEWQIVTPPAGLKAATSYITGLTENQEKIRIYAM 15337  
QY 168 ----- 167  
Db 15338 NSEGLGEPALVPCTPKAEDRMLPPEI ELADLRKVVTIRACCTLRLFVPIKGRPDPEVKW 15397  
QY 168 ----- 167  
Db 15398 ARDHGESLDKASIESASSYTLIVGNVRFDSGKYILTVENSSGSKSAFVNVRLDTPGP 15457  
QY 168 ----- 167  
Db 15458 PODLKVKVTKSVTLTWDPPLLDGSGSKKNYIVEKRESTRKAYSTVATNCHKTSWKVDO 15517  
QY 168 ----- 167



Db 11138 PVINYIVEKQTRKDTGWVSSGSKTKLIPHLOKGCYVFRVRAENKIGVGPLDSTP 11197  
QY 154 ----- 153  
Db 11198 TVAKHKFSPSPGKPVVTDITENAAVSWTLPKSDGSGPITGYMYERREVTGKWVRNK 11257  
QY 154 ----- 153  
Db 11258 TPIADLKFRVTGLYEGNTYEFVRFAENLAGLSKPSSPDIKACRPKPPGPPINPKLKD 11317  
QY 154 ----- 153  
Db 11318 KSRETADLVWTKPLSDGSGPILGYVVECOKPGTAQWNRINKDELIRQCAFVPGLIEGNE 11377  
QY 154 ----- 153  
Db 11378 YRFRIKAANIVGEGEPRELAESVIAKDILHPPEVELDVTCDVITVRVQGTIRILARVG 11437  
QY 154 ----- 153  
Db 11438 RPEPDITWTKGKVLVRERKRVLDIOLPRVELQIKEAVRADHGKYIISAKSSGHAQGS 11497  
QY 154 ----- 153  
Db 11498 IVNVLDROPQCNLKVNTYKENCITISWENPLDNGGSEITNFIVEYRKNQKGSIVASD 11557  
QY 154 ----- 153  
Db 11558 VTKRLIKANLANNEYFRVCAENKVGVTIETKTPIILAINIDRPGEPENLHIADKCK 11617  
QY 154 ----- 153  
Db 11618 TFYLKWRRPDYDGGSPNLSYHVERRLKGGDDWVRHKGSIKETHYMDRCVENQIYEF 11677  
QY 154 ----- 153  
Db 11678 VQTKNEGGSWVKEEVVVKEDLQKPVLDLKLSGVLTWKAGDTRLEAGVRGKPPPEVA 11737  
QY 154 ----- 153  
Db 11738 WTKDKDATLITRSPRVKIDTRADSKFSLTAKRSDGKGYVVTATNTAGSFVAYATVNV 11797  
QY 154 ----- 153  
Db 11798 DKPGPVRLKIVDVSSDRCTVCDPDDGGEIQNYILEKCTKRWVWSTYSATVLTGP 11857  
QY 154 ----- 153  
Db 11858 TTVTRLIEGNEYIFRVAENKIGTGPPTBSKPVIAKTKYDKFGRPDPEVTVKVSKEEMTV 11917  
QY 154 ----- 153  
Db 11918 VNNPEYDGGKSTGYFLEKKEKHSRWVPVNSAIPERRMKVQNLPLDHEYQFRVKAEN 11977  
QY 154 ----- 153  
Db 11978 EIGIPELPSRPPVAKDIEPPGPPTNFRVVDTTKHSITLWGKPVYDGGAPIIGYVVE 12037  
QY 154 ----- 153  
Db 12038 MRPKIADSPDEGWKRCNAAAQLVRKEFTVTSLDENQEYFRVCAQNVGIGRPAELKEA 12097  
QY 154 ----- 153  
Db 12098 IKPKEILEPPEIDLASMRKLVIRAGCPILFAIVRGPAPKVTRWKVIGIDNVVRKGOV 12157  
QY 154 ----- 153  
Db 12158 DLVDTMAFLVIPNSTRDDSGKYSLTLPNAGEKAVFNVVRVLDTPGVSDLKVSDVTKTS 12217  
QY 154 ----- 153  
Db 12218 CHVSWAPPENDGGSQVTHYIVEKREADRKTWSTVTPVEVKKTSFHVNLNLPVGNEYFRVFTA 12277

QY 154 ----- 153  
Db 12278 VNEYGPGVTDVPKPVLASDPLSEDPDPKRLKATLWMLPPLRDGGAIDGVI 12337  
QY 154 ----- 153  
Db 12338 ISYREEQPADRWTEYSVVKDLVSVVTGLKEGKYYKFRVAARNAVGSUPREAEGYEAK 12397  
QY 154 ----- 153  
Db 12398 EQLLPKILMPEQIITIKAGKRLIEAHVYKPKHPCTCKWKKEDEVVTTSSHAVHRAADSS 12457  
QY 154 ----- 153  
Db 12458 ILI IKDVTBKDSGYSLTAENSSGTDQKIKVVMADAPGPPPPFDISDADACSLSWH 12517  
QY 154 ----- 153  
Db 12518 IPLEDGGSNITNYIVEKCDVSRGDWWTALASVTKTSCRVGKLIPGOEYIFRVAENRFGI 12577  
QY 154 ----- 153  
Db 12578 SEPLTSPKMAQFPFGVPSEPKNARVTKVKNKDCIFVAVDRPDSGSPILIGYLIERKRN 12637  
QY 154 ----- 153  
Db 12638 SLLWKANDTLVRSTEYPCAGLVEGLEYSFRIVALNKAAGSSPPSKPTEYVTARMPVDPGP 12697  
QY 154 ----- 153  
Db 12698 KPEVIDVTKSTVLIWARKPHDGGSKIIGYFVEACKLPDGKWRCNTAPHQIPQEEYTAT 12757  
QY 154 ----- 153  
Db 12758 GLEBRAQYQFRAIARTAVNISPSESDPVTILAENVPDRIDLSVAMKSLTIVKAGTNVC 12817  
QY 154 ----- 153  
Db 12818 LDATVFGKPMPTVSMKKGDTLLKPAEGIKMAQORNLCTLEFSVNRKSDGDTITAENSS 12877  
QY 154 ----- 153  
Db 12878 GSKSATIKLVLDKFGPPASVKINKMYSDRAMLSWEPPLDGGSEITNYIVDKRETSRP 12937  
QY 154 ----- 153  
Db 12938 WAQVSATVPITSCSVEKLLIEGHEYQFICAENKYGVGDVPFTEPAIAKNPYDPPGRCDPP 12997  
QY 154 ----- 153  
Db 12998 VISNITKDHMTYSWKPPADDGGSPITGYLLEKRETOAVNWKVNRKPIIERTLKATGLQE 13057  
QY 154 ----- 153  
Db 13058 GTEYEFRTAINKAGPKPSDASKAAYARDPOYPAPPAPKVYDITRSSLVSLWSGKPAY 13117  
QY 154 ----- 153  
Db 13118 DGGSPILGYLVEVKRADSDNWRVRCNLQNLQKTRFEVTGLMEDTQYQFRVAVNKGIGSD 13177  
QY 154 ----- 160  
Db 13178 PSDVPDKHYPKDILLIPPEGEHDADLRKTLILRAGVTMRLYVPVKGRPPPKITWSKPNVNL 13237  
QY 161 ----- 160  
Db 13238 RDRIGLDIKSTDFDPLRCENVNKYDAGKYILTLENSCGKKEYTIVVKVLDTPGPPINVT 13297  
QY 161 ----- 160  
Db 13298 VKEISKDSAYTWEPPIIDGGSPINNYVQKRDARERKSWSTVTECSKTSFRVNPNEEGK 13357

QY 133 ----- 132  
Db 8978 YTVASNRLGVSFRNVHVEYDRPSPRNLA VTDIKAESCYLTWDAPLNDGSGSEITHYVI 9037  
QY 133 ----- 132  
Db 9038 DKRDASRKAEBEVTNTAVEKRYGIWKLI PNQYEFVRVANKYGISDECKDKVVIQD 9097  
QY 133 ----- 132  
Db 9098 PYRLPGPPGPKVLARTKGSMLVSWTPPLDNGGSPITGYWLEKREGE SPYWSRVSRAPIT 9157  
QY 133 ----- 132  
Db 9158 KVG LGVEFNVPRLLGKVKQFRAMAINAAGIGPPSEPSDEVAGDPIFPFGPPSCPEVK 9217  
QY 133 -----PPSRD----- 137  
| : : :  
Db 9218 DKTSSISLGWKPPAKDGGSPKIGYIVEMOE G TDMKRVNPD KLIITTCCECVVNLKEL 9277  
QY 138 ----- 137  
Db 9278 RYFRFRVAVNEAGESEPSDTTGEIPATDIQEPEV FIDIGA QDCLVC KAGSIRIPAVI 9337  
QY 138 ----- 137  
Db 9338 KGRPTPKSSWFDGAKKAMKDGVDHIDIPEDAQIETAENSSV I I I PECKRSHTGKYSITAK 9397  
QY 138 ----- 137  
Db 9398 NKAGQKTANCRVKMDVPGPPKDKLVSDITRGSCRLSWKMPDDGGDR IKG YVIEKRT ID 9457  
QY 138 ----- 137  
Db 9458 GKAWTKVNPCGSTTFVVPDLLSEQYFFRVRAENRFGIPGPVETIQRTTARDPIYPPDP 9517  
QY 138 ----- 137  
Db 9518 PIKLGITLITKNVHLSWPKPKNDGSGPVTHYIVECLANDPTGT KKEAWRQCNKRDVEEL 9577  
QY 138 ----- 137  
Db 9578 QFTVEDLVEGGEYFRVKAVNAAGSVKPSATVGCQCRDMPDSIDLKEFMEVEBGTN 9637  
QY 138 ----- 137  
Db 9638 NIYAKIKGVFPFPLTWFKAPPKKPDNKEPVL YDTHVNKL VVDDTCTLVIPQSRSDTGLY 9697  
QY 138 ----- 137  
Db 9698 TITAVNLTGASKEMRLNVLGRGPPVGP I K FESVADQMTLSWFPKDDGSGKIITNYVI 9757  
QY 138 ----- 137  
Db 9758 EKREANRKTWVVSSEPFECTYITPKLLEGHEYVFRIMAQNKYIGEP L DSEPETARNLF 9817  
QY 138 ----- 137  
Db 9818 SVPGAPDKPTVSSVTRNSMTVNWEEPEYDGGSPVTGYWLEMKD TTSKRWRVNRDPIKAM 9877  
QY 138 ----- 137  
Db 9878 TLGVS YKVTGLIEGSDYQFRVYAINAAGVPASLPSPDPATARDPIAPGPPFPKVTDWTK 9937  
QY 138 ----- 137  
Db 9938 SSADLEWSPLKDGSKVTGYIYEYKEEGKEWEKGRKEVGT KLVVTGLKEGAFYKFR 9997  
QY 138 ----- 137  
Db 9998 VSAVNIAGICEPCEVTDVIEKDRLVSPDLQLDASVRDRIVVHAGGVIRIIAIVSGKPPP 10057  
QY 138 ----- 137

Db 10058 TVTWNMNER TLPQEA T IET TAISSSMV I KNCQRSHOGVSYLLAKNEAGERKKTIIVDVLD 10117  
QY 138 ----- 137  
Db 10118 VPGPVGT PFLAHLNTNESCKLTWFSPEDDGGSPITNYVIEKRESDRRAWTPVYTVYTRQN 10177  
QY 138 -----ELTKNQVS 145  
| : : :  
Db 10178 ATVOGLIQKAYFFRIAAENSICGPFVETSEALVIREPITVPERPEDLEVKEVTKNTVT 10237  
QY 146 LT----- 147  
| : : :  
Db 10238 LTWNPKYDGGSEIINVLESRLIGTEKPHKVTNDNLLSRKYTVKGLKEGDTYEYRSVAV 10297  
QY 148 ----- 147  
Db 10298 NIVGQKPSFCTKPI TC KDELAPPTLHLDFRDKLTIRVGEAFALTGRYSGKPKVSWFK 10357  
QY 148 -----CLV-----KGF----- 153  
| : : :  
Db 10358 DEADVLEDDRTHIKTTPATLAL EKI KAKRSDSGKYCVVVENSTGSRKGFQCVNVVDHJPG 10417  
QY 154 ----- 153  
Db 10418 PVGPVSFDEVT KDMYVISMKPP LDDGSGKITNYIIEKKEVGDVWMPVTSASAKTCKVS 10477  
QY 154 ----- 153  
Db 10478 KLEKGDYIFRIHAENLYGISDPLVSDSMKAKDRFRVPDAPDQPIVTEVTKDSALVTWK 10537  
QY 154 ----- 153  
Db 10538 PHDGGKPI TNYIIEKRETMSKRWARTKDIHPYTFKRPD LLEGGCOYEPFRVAENBEGI 10597  
QY 154 ----- 153  
Db 10598 GDPSPSKPVFAKDIAPKPSPPVNP EALDTTCNSVDLTWQPPRRHDGSGKITGVIYEQV 10657  
QY 154 ----- 153  
Db 10658 GDEWRANHTPESCPETRYKVTGLRDGQTYKFRVLAVNAAGESDPAHVPEPVLVKDRLE 10717  
QY 154 ----- 153  
Db 10718 PPELIIL DANWAREQH I KVGDTLRLSALIKVPPFKVTWKEDRDAPT KARIDVTPVGSKL 10777  
QY 154 ----- 153  
Db 10778 EIRNAAHEDGGIYSLTVENPAGSKTVSVKVLVDKPGPPRDLEVSEIRKDCYLTWKEPL 10837  
QY 154 ----- 153  
Db 10838 DGGSVITNYVVERRDVASAQNSPLSATS KKKSHFAKHLNEGNOYLFRVAENQYGRGPF 10897  
QY 154 ----- 153  
Db 10898 VETPKPIKALDPLHPGPKDLHHVDVDKTEVSLVWNKPD RDGSGSPITGYLVEYQEGTQ 10957  
QY 154 ----- 153  
Db 10958 DWIKFKVTNLECVVTGLOOGKTYRFRVKAENIVGLGLPDTTIP IECQEKLVPPSVELDV 11017  
QY 154 ----- 153  
Db 11018 KLIEGLVWAGTTVRPPAIRGVPTAKWTTDGS EIKTDEHYTVETDNFSSVLTKNCL 11077  
QY 154 ----- 153  
Db 11078 RRDTEYQITVSNAGSKTVAVHLTVLDVPGPPPTGPI NILDVTPEHMTISWQPPKDDGGS 11137  
QY 154 ----- 153

Db	6758	VVEGOYSFRVRAQNRIGVGKPSAATPEVKVADPIERPSPVNLTSDDQTOSSVOLKWE	6817
QY	133	-----	132
Db	6818	PLKGGSPILGYIIERCEEGKDNWIRCNMKLVPETYKVTGLEKGNKYLRVSAENKAGV	6877
QY	133	-----	132
Db	6878	SDPSELLPLTADDAFVEPTMDLSAFKDGLEIVIVPNPITILVPSTGYPRPTATWCFGKV	6937
QY	133	-----	132
Db	6938	LETGDRVMKTLTSAELVISPERSDKGIWTLKLENRKVTISGEIDVNVIAEPSAPKEL	6997
QY	133	-----	132
Db	6998	KFGDITKDSVHLTWEPDDGGSPLTGYVVEKREVSRTWTVMDFVTDLEFTVPDLVQG	7057
QY	133	-----	132
Db	7058	KEYLFKVCARNKCGEPAYVDEPVNMSTPATVPDPENVKWRDRTANSIFLTWDPKND	7117
QY	133	-----	132
Db	7118	GGSHIKGYIVERCPRGSDKWVACGEPVAETKMEVTGLEGKWYAYRVKTLNRQASKPSR	7177
QY	133	-----	132
Db	7178	PTBEOAVDQEAPEIFDLVKKLAGLTVKAGTKIELPATVTVGKPEPKITWTKADMILKOD	7237
QY	133	-----	132
Db	7238	KRITIENVPKKSTVTIVDSKRSDTGTVYIEAVNVCGRATAVEVNVLDKPGPAAFDITD	7297
QY	133	-----	132
Db	7298	VTNESCLLTWNPRDDGGSKITNYVVERRATDSEVWHKLSSTVKDTNFKATKLIPKEYI	7357
QY	133	-----	132
Db	7358	FRVAENWYGAGEPVQASPIYAKYQDPDPPGPPTRLEPSDITKDAVLTWCEPDGGSPI	7417
QY	133	-----	132
Db	7418	TGYWVERLDPDQKWRCKNMPVKDTYRVKGLTNKKYRFRVLAENLAGPGKPSKSTEP	7477
QY	133	-----	132
Db	7478	ILIKDPIIDPPWPGKPTVKDVGKTSVRLAWTKPEHGGAKIESYVIEMLKTGTDEWVRVA	7537
QY	133	-----	132
Db	7538	EGVPTTQHLLPGLMEGOEYSFRVRAVNKAGESEPSDPSDVLCREKLYPPSPRWLEVIN	7597
QY	133	-----	132
Db	7598	ITKNTADLKWTVPEKGGSPITNYIVEKRDVRKGMQVDTTVTKDTKCTVPTLTEGSLYV	7657
QY	133	-----	132
Db	7658	FRVAENAIQSDYTEIEDSVLAKDTFTTGPYPALAVVDVTKRHVDLKWPEPKNDGGRP	7717
QY	133	-----	132
Db	7718	IQRVYIEKKERLGRWVKAGTAGPCNFRVTDVTEGTEVQFVRAENAGVCHPSEPT	7777
QY	133	-----	132
Db	7778	ILSIEDPTSPSPPLDLHVTAGRKHIAIAWKPPKNGSGPIIGHVEMCPVGTCKWVRV	7837
QY	133	-----	132
Db	7838	NSRPIKDLKFKEGVVPOKEYVLRVRAVNAIGVSEPSISENVVAKDPDCKPTIDLETH	7897

QY	133	-----	132
Db	7898	DIIVIEGEKLSIPVFFRAVPPTVSWHKDGKEVKASDRLTMKNDHISAHLEVPKSVRADA	7957
QY	133	-----	132
Db	7958	GIYITITLENKLSATASINVKVIGLPGCKDIKASDITKSSCKLTWBPFPDGGTPIIHY	8017
QY	133	-----	132
Db	8018	VLEREAGRRTYIPVMSGENKLSWTVKDLIPNGEYFFRVRANKVVGGEYIELKNPVAQ	8077
QY	133	-----	132
Db	8078	DPKOPDPDPDVEVHNPTAEAMTITWKPPLYDGGSKIMGYIIIEKIAKEERWKRNEHLV	8137
QY	133	-----	132
Db	8138	PILTYTAKGLEEGKEYQFRVRAENAGISEPSRATPPTKAVDPIADPKVILRTSLEVRG	8197
QY	133	-----	132
Db	8198	DEIATDASISGSPYPTITWIKDENVIYPEIEKKRAAPLVRRRRKEVQEEFFVLPLTQRL	8257
QY	133	-----	132
Db	8258	SIDSKKESQLRVRDSLRPDHGLYMIKVENDHGIAKAPCTSVLDTPGPPINFVPEDIR	8317
QY	133	-----	132
Db	8318	KTSVLCKWEPLDDGGSEIINYITLEKKDKTKPDSEINVTSTLRHCKYSVTKIEGKEYL	8377
QY	133	-----	132
Db	8378	FRVRAENRFGPGPCVSKPLVAKDPGPPADPKPIVEDVTSNSMLVKNWKNPKDNGSPIL	8437
QY	133	-----	132
Db	8438	GYWLEKREVNSTHWSRVNKSLLNALKANVDGLLEGLTYVPRVCAENAGPGKFPSPDPK	8497
QY	133	-----	132
Db	8498	TAHDPISPPGPPIPRVTDTSSTTIELEWEPAPFNGGGEIVGYFVDKQLVGTNKNWCRTEK	8557
QY	133	-----	132
Db	8558	MIKROYTVKEIREGADYKLRVSANVNAAGBPGPETQPTVAEPQEPAPVELDVSKGGI	8617
QY	133	-----	132
Db	8618	QIMAGKTLRIPAVVTGRPVPTKVWTKEEGELDKDRVIDNVGTKSELLIKDALRKDHGRY	8677
QY	133	-----	132
Db	8678	VITATNSCGSKFAARVEVDVPGPVLDPVVVTRNKKMLLNWSDPEDDGGSEITGFIIE	8737
QY	133	-----	132
Db	8738	RKDAKMTWRQPIETERSKCDITGLLEGOEYKFRVIAKNKFGCGPPVEIGPILAVDPLGP	8797
QY	133	-----	132
Db	8798	PTSPERLTYTERQSTITLDWKEPRNGGSPIQGYIIIEKRRHDKPDFERVNKRCLCPTTSF	8857
QY	133	-----	132
Db	8858	LIVENLDEHOMYEFVRKANVEIGESEPSLPNVVQDEVPPTIKLRLSVRGDTIKVKAGE	8917
QY	133	-----	132
Db	8918	PVHIIPADVTGLPMPKIENSKNETVIEKPTDALQITKEEVSRSSEAKTELSIPKAVREDKCT	8977

Qy	115	-----	114
Db	4598	SGGKPPDEAPFTYQKAVPLKFVKEIKDIILTESEFVGSSAIFECVSPSTAITWMKO	4657
Qy	115	-----	114
Db	4658	GSNIRESPKHFRIADGKORKLHIIDVQLSDAGEYTCVLRGNKEKTSTAKLVVEELPVRF	4717
Qy	115	-KTIS- 	118
Db	4718	VKTLEEEVTVVKQPLYLSCELNKERDVVRKDKIVVEKPGRIVPGVIGLMRALTINDA	4777
Qy	119	-----	118
Db	4778	DDTAGTYTVVENANNLECCSVKVEVIRDWLVPKPIRQHVKPGTAIFACDIAKTP	4837
Qy	119	-----	118
Db	4838	NIKFKGYDEIPAEPNDKTEILRDGNHLYLKIKNAMPEDIAEYAVEIEGKRYPAKLTLGE	4897
Qy	119	-----	118
Db	4898	REVELLKPIEDVTIYEKESADAEISEADIPGQWKLKELLRPSPTCEIKAEKGKRLT	4957
Qy	119	-----	118
Db	4958	LHKVKDQAGEVLYQALNAITTAILTVKEIELDFAVPLKDVTPERRQARFECVLTREAN	5017
Qy	119	-----	118
Db	5018	VIWSKGPDIIKSSDKFDIIADGKKHILVINDSQFDDGCVYTAEEVGKTSARLFTVGTGIRL	5077
Qy	119	-----	118
Db	5078	KFMSPLEDOTVKEGETATFVCELSHEKMHVWVFNDAKLHSTRTVLISSEKTHKLENKE	5137
Qy	119	-----	118
Db	5138	VTLDDISQIAQVKELSTAQLKLEADPYFTVKLHDKTAVKDEITLKCEVSKDVPVKW	5197
Qy	119	-----	118
Db	5198	KDGEIIVPSPKYSIKADGLRILKIKADLKRGYVDCGTDKTKANVTVEARLIEVE	5257
Qy	119	-----KAGOP- 	124
Db	5258	KPLYGVEFVGETAHFEIELSEPDVHGQWKLKGOPLTASPCDCEIIEDGKKHILHNCQL	5317
Qy	125	-----REPO- 	128
Db	5318	GMTGEVSFOAANAKSAANLKVKELPLFITPLSDVKVFKEDEAKFECEVSREPKTFRWLK	5377
Qy	129	-----	128
Db	5378	GTOBITGDRPELIKDGTKHSWIKSAAFDEAKYMFEEADKHTSGKLIIEGIRLKFLTP	5437
Qy	129	-----	128
Db	5438	LKDVTAKEKESAVTVVELSHDNIRVKWPKNDQRLHTRTSVSMQDEGKTHSTTFKDLSDID	5497
Qy	129	-----	128
Db	5498	TSQIRVEAMGMSSEAKLTVLEGDPHYFTGKLQDYGVEKDEVILQCEISKADAPVKWFEDG	5557
Qy	129	-----	128
Db	5558	KEIKPSKNAVIKTDGKKRMLILKALKSDIGQYTCDCGTDKTSKGLDIEDREIKLVRPLH	5617
Qy	129	-----	128
Db	5618	SVEVNETARFETEISEDDIHANWKLKGEALLQTPDCEIKEGKIHSVLHNCRLDQTG	5677
Qy	129	-----	128

Db	5678	GVDFQOANVKSSAHLRVKPRVIGLLRPLKDVTVTAGEATATEDCELSYEDIPVWYLKGGK	5737
Qy	129	-----	128
Db	5738	LEPSDKVVPRSEGVHTLTLRDVKLEDAGEVOLFAKDPKTHANLFVKEPPEFTKPLEDDQ	5797
Qy	129	-----	128
Db	5798	TVEGATAVLECEYSRENAKVKWFKNGTEILSKKYEIVADGRVPRKLVHDCPTPEDIKTY	5857
Qy	129	-----	128
Db	5858	TCDAKFKTSCNLNVPPHVEFLRPLTDLQVREKEMARFECELSRENAKVKWFKDGAIEIK	5917
Qy	129	-----	128
Db	5918	KGKYYDIISGAVRILVINKLCLDDEAEYSCEVRTARTSGMLTVLEBAVFTKNLANIEV	5977
Qy	129	-----	128
Db	5978	SETDTIKLVCEVSKPGAIEVHWKGDEEIIETGRYEILTEGRKRILVIONAHLEDAGNVC	6037
Qy	129	-----	128
Db	6038	RLPSSRTDGKVKVHELAAEFISKQNLLEIGEKAEFVCSISKESFPVQWKRDDKTLESG	6097
Qy	129	-----	128
Db	6098	DYDVIADGKKRVLVVKDATLQDMGTYYVMVGAARAAHLTVIEKLRIVVPLKDRVKEQ	6157
Qy	129	-----VYTL- 	132
Db	6158	QEVVFNCEVTEGAKAKWFRNEBAIFDSSKYIILOKDLVYTLRIRDAHLDDOQANYVSLT	6217
Qy	133	-----	132
Db	6218	NHRGENVKSANLIVEEEDLRIVEPLKDIEETMEKKSVTFWCKVNLNVTLAKTKNGBEVP	6277
Qy	133	-----	132
Db	6278	FDNRSYRVDRYKHKMLTIKDCGFPDEGEYIVTAGODKSVABELLITEAPTEFVEHLEDQTV	6337
Qy	133	-----	132
Db	6338	TEPDDAVFSQOLSREKANVKWYRNGREIKEGKYKFEKDGSIHRLIINKDCRLDDECEYAC	6397
Qy	133	-----	132
Db	6398	GVEDKSRARLFVEEIPVEIIRPPQDILEAPGADVFLAELNKDKVEQWLRNNWVVQG	6457
Qy	133	-----	132
Db	6458	DKHOMSEGIHRLQICDIKPRDQGEYRFIAKDKEARAKLEAAAPKIKTADQLVDVG	6517
Qy	133	-----	132
Db	6518	KPLTMVVPYDAYPKAAEWFKENEPLSTKTTIDTTAOTSFRILEAKKGDKGRYKIVLQNK	6577
Qy	133	-----	132
Db	6578	HGKAEGFTNLKVIDVPGVPRNLEVTEFDGCVSLAWBEPLTDGSGKIIGYVVERRDKRK	6637
Qy	133	-----	132
Db	6638	TWVLATDRASCEFTVTGLQGGVEYLFVRSARNRVGTGEPVETDNPVEARSKYDVPGP	6697
Qy	133	-----	132
Db	6698	LNVITDVNRFGVSLTWEPPEYDGGAEITNVYIELRDKTSIRWDTAMTVRAEDLSATVTD	6757
Qy	133	-----	132

Db 2378 NYSFTIPALGLSTGRVSYSVDVITPLKDNVNIIEGTAKVLECKVSPDVTSVKWLINDE 2437  
QY 74 ----- 73  
Db 2438 QIKPDRQVAIVGTPKQRLVINRTHASDEGPKLIVGRVETNCNLSVEKIKIIRGLRDLT 2497  
QY 74 ----- 73  
Db 2498 CTETQNVVEVELSHSGIDVLNFKDKIEKPSKYYKIEAHGKIYKLTVLNMMKDDGKYT 2557  
QY 74 ----- 73  
Db 2558 FYAGENWTSGLTVAGGAISPLTDQTVAESQEAUFCEVEANPDSKGWLRDGHPLPTN 2617  
QY 74 ----- 73  
Db 2618 NIRSSEDGHRLLIATAATKDDDIGEYTYKVATSKTSAKLKVEAVKIKTKLNLTVETQD 2677  
QY 74 ----- 73  
Db 2678 AVFTVELTHPNVKGWIKNGVLESNEKYAISVKGTIYSLRIKNCIAIVDESIVYGRGLGR 2737  
QY 74 ----- 73  
Db 2738 LGASARLHVETVKIITKKPKDVTALENATVAFEVSVSHDTPVKWPHKSVETKPSDKHRLV 2797  
QY 74 ----- 73  
Db 2798 SERKVHKLMLQNLSPSDAGEYTAVVQLECKAKLFTVETLHITKTMKNIEVPETKTASFEC 2857  
QY 74 ----- 73  
Db 2858 EVSHFNVPMSMLKNGVEIEMSEKIVVQGLHQLHIMNTSTEDSAEYTFVCGNDQVSAT 2917  
QY 74 ----- 73  
Db 2918 LTVTPIMITSMLKDINAEKDTITFEVTVNVEGISYKWLKNGVEIKSTDKCOMTKKLTH 2977  
QY 74 ----- 73  
Db 2978 SLNIRNVHFGDAADYTFVAGKATSTATLYVEARHIEFRKHIDIKVLEKKRAMFECEYSE 3037  
QY 74 ----- 73  
Db 3038 PDITVQWKKDDQLOITDRIKIQEKYVHRLIIPSTRSDAGKTYTVAGNVNSTAKLVE 3097  
QY 74 ----- 73  
Db 3098 GRDVRIRSIKKEVQVIEKQRAVVEFEVNEDDVDHAWYKDGIEINFQVQERHKYVVERRIH 3157  
QY 74 ----- 73  
Db 3158 RMFISETRQSDAGEYTFVAGNRSSVTLYVNAPEPQVLQELQPVTVQSGKPARFCAMIS 3217  
QY 74 ----- 75  
Db 3218 GRPQPKISWYKEEQLLSTGFKFKFLHDCQEYTLILLIEAFEDDAAVTCEAKNDYGVATTS 3277  
QY 76 ----- 75  
Db 3278 ASLSVEVPEVSPQEMPVYPPAIITPLQDVTVTSBGPARFCQRCVSGTDLKVSWSKDKK 3337  
QY 76 ----- 83  
Db 3338 IKPSRFRMTQFEDTYOLEIAEAVPEDEGTFTFVANNVAVQVSSSTANLSLEAPESILHER 3397  
QY 84 ----- 83  
Db 3398 IEQEIEMEMKEFSSFLSAEEGLHSAELQLSKINETLELLESPPVPTKFDSEKEGTGP 3457  
QY 84 ----- 87  
Db 3458 IFIKEVSNADISMGDVATLSVTIGIPKPKIQWPFNGVLLTPSADYKFVFDGDDHSLIIL 3517

QY 88 ----- 87  
Db 3518 FTKLEDEGEYTCMASNDYKKTICSAYLKLNSKGECHKDTETESAVAKSLEKLGPCPHF 3577  
QY 88 ----- 91  
Db 3578 LKELPIRCAOGLPAIFEYTVVGEPAFTVTFKFNKQLCTSVYYTIIHNPNGSTGTFIVND 3637  
QY 92 ----- 91  
Db 3638 POREDSGLYICKAENMLGESTCAELLVLLDPTDTPCKAKSTPEAPEDFPQPLKGP 3697  
QY 92 ----- 91  
Db 3698 AVEALDSQEIATFVKDTILKAALITEENQQLSYEHIAKANELSSOLPLGAQELQILBQ 3757  
QY 92 ----- 91  
Db 3758 DKLTPESTREFLCINGSIHFOPLKEPSPNLQIIVOSQKTFSEKIGILMPEPEQAVLSD 3817  
QY 92 ----- 91  
Db 3818 TEKIFPSAMSTEQINSLTVEPLKTLAEPEGNYPOSSIEPPMHSHYLTSVAEVLSLKEKT 3877  
QY 92 ----- 91  
Db 3878 VSDTNREQRVTLQKQEAQSALILSQSLAEGHVESLQSPDVMISQVNYEPLVPSEHSECTEG 3937  
QY 92 ----- 91  
Db 3938 GKILIESANPLENAGDSAVRIEKGSLRFPALAEKQVLLKEHSDNVMPDQIIEK 3997  
QY 92 ----- 95  
Db 3998 REPVAIKKQVQVQGRDILLKESLLSGIPEEQRNLKIQICRALQAAVASEQGLFSEWL 4057  
QY 96 ----- 95  
Db 4058 NIEKVEEAVNITQPRHIMCMVLTSAKSVTEVTIIIEDVDPQAMKLMELDALCAI 4117  
QY 96 ----- 95  
Db 4118 IYEEIDILTAEGPRIQOQAKTSLQEEMDSFGSQKVEPITEPEVESKYLISTEEVSFNV 4177  
QY 96 ----- 95  
Db 4178 QSRVKYLDATPVTKGVASAVSDEKQDESLEKPSKEEKSSESGTEEVATVKIQAEGL 4237  
QY 96 ----- 95  
Db 4238 IKEDGPMIHTPLVDTVSEGDIVHLTTSITNAKEVNWYFENKLVPSDEKFKCLODQNTYT 4297  
QY 96 ----- 98  
Db 4298 LVIDKVNTEHDHGEYVCEALNDSGATATSAKLTVVKRAAPVYKRIEPLVLEVALHAKFT 4357  
QY 99 ----- 106  
Db 4358 CEIQSAPNVRFQWFKAGRIEYEDSKSIRSSKYISSLELTRQVVVDCGEYTCCKASNEYS 4417  
QY 107 ----- 106  
Db 4418 VSTATLATVTVPGKEKVKRLLPERKBPKEEVKLSVLRKRPEEPKVEPKLEKVK 4477  
QY 107 ----- 106  
Db 4478 PAVPEPPPPKPVEEVEVTVTKRERKIPEPTKVPEIKPAIPLPAPKPKPAEVKTKP 4537  
QY 107 ----- 114  
Db 4538 PPVEPTPIAAPVTVVVGKKAEPKAAKPKGPIKGVPKKTPSPIEAERKLRPG 4597

QY 10 ----- 9  
Db 218 QTRIEKIEAHFARSATVEMVIDGAAGQOLPHKTPRIPPKPSRSPTPPSIAAKAQL 277  
QY 10 ----- 9  
Db 278 AROQSPSPIRHSPSPVRHVRAPTPSPVRSVSPAARISTSPRSPLLMRKTOASTVAT 337  
QY 10 ----- 9  
Db 338 GPEVPPKQEGYVASSSEAEEMRETTLTSTQINTEBEREGRYGVQOVQVTSIGAAGAAAS 397  
QY 10 ----- 9  
Db 398 VSASASYAAEAVATGAKEVKQDADKSAAVATVVAADVMAVRREPVISAVQTAQRTTTA 457  
QY 10 ----- 9  
Db 458 VHIQPAQEQVRKEAKTAVTKVVVVAADKAKEQELKSRKEIITTKQBQMHVTHEQIRKET 517  
QY 10 ----- 9  
Db 518 EKTFFVKVLSAAKAKEQETRISEEITKKQOVTOEALMKETRTKVVPKVIVATPKVKEQ 577  
QY 10 ----- 9  
Db 578 DLVSRGREGITTKREQVOITQEKMRKEAKETALSTIAVATAKAKEQETILTRTETMATRQ 637  
QY 10 ----- 9  
Db 638 EQIQVTHGKVDGKKAABAVATVVAADVQARVREPREGHLEESYAOQTTLLEYGYKERISA 697  
QY 10 ----- 9  
Db 698 AKVAEPPQRPASBPHVVPKAVPRVIOAPSETHIKITTDQGMHISOIKKTTDLTTERLV 757  
QY 10 ----- 9  
Db 758 HVDKRPRTASPHFTVSKISVPKTEHGYEASIAATLQKELSATSQAQKITKSVKAPT 817  
QY 10 ----- 9  
Db 818 VKPSETRVRAEPTPLPOFPFADPTDYKSEAGVEVKEVGVSTGTTVREERFEVLHGRE 877  
QY 10 -----CPAP-----ELLGSPSVFL-----PP----- 25  
Db 878 AKVTETARVPAPVEIPVPTPLVSLGNVTVIEGESVTLCHISGYPSPVTWYREDYQI 937  
QY 26 ----- 25  
Db 938 ESSIDFQITFQSGIARLMIREFAFEDSGRFTCSAVNEAGTVSTCYLAVQVSEFEKETT 997  
QY 26 ----- 25  
Db 998 AVTEKFTTEKRFVESRDVMTDSTLSEQAQGPGEPAAPYFITKPVVQKLVEGGSVVFQC 1057  
QY 26 -----PKP----- 28  
Db 1058 QVGNPKPHYWKSGVPLTTGYRYKVSYNKQTECKLVISMFTFADDAAGEYTIIVRNKHG 1117  
QY 29 ----- 28  
Db 1118 ETSASALLEADYELIMKSQEMLYQTQVTAQVPEVGETAPGFVSEYEKEQA 1177  
QY 29 -----KDTLM----- 33  
Db 1178 LIRKMAKDTVVVTVTVEDQEFHISSEERLIKEIERYIKTTLLELLEDGEKMAVDI 1237  
QY 34 ----- 33  
Db 1238 SESEAVESGFDLRKNRILEGMVTFHCKMSGYPLPKIANYKDGRKIKHGERYQMDFLQ 1297  
QY 34 ----- 33

Db 1298 DGRASLRIPVVLPEDEGIYTAFAASNIKGAICSGKLYVEPAAPLGAPTYIPTLEPVSRI 1357  
QY 34 -----ISRTP----- 38  
Db 1358 SLSPRSYSRSPIRMSPARMSPARMSPARMSPARMSPARMSPARMSPARMSPARMSPARMS 1417  
QY 39 ----- 38  
Db 1418 VSFKLEGANCNRPDLKVVGPRMPETFWFDGQOQIVNDYTHKVVIKEDGTOSLIIVPATPS 1477  
QY 39 ----- 38  
Db 1478 DSGEWTVAQNAGRSSISVILTVEAVEHQVPMFEVKLNKVNKEGSRLEMKVATGNP 1537  
QY 39 ----- 38  
Db 1538 NPDIVMLKNSDIIIVPHKYPKIRIEGTKEAALKIDSTVSQDSAWYTATAINKAGRDTRC 1597  
QY 39 ----- 38  
Db 1598 KVNVEVEFAPEPERKLIIPRGTYRAKEIAAPELEPLHLRYGQEQWEGDLYDKKQOKP 1657  
QY 39 ----- 38  
Db 1658 FFKKLTSLRKRFGPAHFECRLTPISDPTMVVWLHDGKPLEAANRLRMINFEGYCSLD 1717  
QY 39 -----EUTC----- 42  
Db 1718 YGVAISRDSGIITCRATNKYGTDHTSATLIVKDEKSLVESQLPEGRKGLQRIEELERMA 1777  
QY 43 ----- 42  
Db 1778 HEGALGVTTDQKEKOKPDIVLYPEVPVLEGETARFCRVCTGYPOPKVNWYNGQLIRK 1837  
QY 43 ----- 42  
Db 1838 SKRFRVRYDGIHYLDIVDCKSYDTGEVKVTAENPEGVIEHKVLEIQOQREDRFSVLRRAP 1897  
QY 43 ----- 42  
Db 1898 EPRPEFHVHPEGKLOFEVQKVRPVDTTETKEVVKLKRAERITHEKVPESEBELRSKFKR 1957  
QY 43 ----- 42  
Db 1958 RTEEGYEAITAVELSKRKDESVEBLLRKTDELLHWTKELTEEEKKALAEKGKITIPT 2017  
QY 43 ----- 42  
Db 2018 FKPDKIELSPSEAPKIFERIQSQTVGQSDAHFRVVRVVGKPDCEWYKNGVKIERSDR 2077  
QY 43 -----VVYDVSHED----- 51  
Db 2078 IYWPEDNVNCELVIIRDVTAEDSASIMWKAINIAGETSSHAFLVLOAKOLITFTQELQDV 2137  
QY 52 -----PEVKFNWYVDGVEVH----- 66  
Db 2138 VAKEDKTMATFECETSEPFVKVYKWKDGMVEVHGEDKYRMHSORVKVHFLSLITLITDSDAED 2197  
QY 67 -----NAKT----- 70  
Db 2198 YSCVLVEDENVKTTAKLIVEGAVVEFVKELQDIEVPESYSGELECI VSPENIEGKWHND 2257  
QY 71 -----KPR----- 73  
Db 2258 VELKSNKYTIITSRGRQNLTVKDVTKEDQGEYSFVIDGKKTCKLKMKPRPRTAILQGLS 2317  
QY 74 ----- 73  
Db 2318 DQKVCBGDIVOLEVKVLSVESVEGVWMDQGEVQPSDRVHIVIDKQSHMLLIEDMTKEDAG 2377  
QY 74 ----- 73



QY 202 ----- 201  
Db 31358 PSQCELEILSISKSVTLQWKEKPCDGGKEILGYWVEYRQSGDSAWKSNKERIKDQF 31417  
QY 202 ----- 201  
Db 31418 TIGGLEATEYEFVFAENETGLSPRRRTAMSIKLTSGEAPGIRKEMKDVTTKLGEAA 31477  
QY 202 ----- 201  
Db 31478 QLSQIOVGRPLPDIKWYRFGKELIOSRYKMWSSDGRTHTLVMTTEEQDEGVYTCIATNE 31537  
QY 202 ----- 201  
Db 31538 VGEVETSSKLLLOATPQHPGYPPLKEKYGAVGSLRLHLVMYIGRPVPAMTWFHCOKLLQ 31597  
QY 202 ----- 204  
Db 31598 NSENITIENTEYTHLVMKNVQRKTHAGKYKVQLSNVFGTDAILDVEIQDKPKTGP 31657  
QY 205 ----- 204  
Db 31658 VIEALLKNSAVISWKPADDGGSWITNVYVEKCEAKGAEWQLVSSAISVTTTCRIVNLTE 31717  
QY 205 ----- 208  
Db 31718 NAGYFVRSQAQTFGSDPLEVSSVLIKSPFEKPGAPGKPTITAVTKDSCVVAWKPPAS 31777  
QY 209 ----- 208  
Db 31778 DGGAKIRNYLEKREKKONKWISVTTEIRETVFSVKNLIEGLEVEYFRVKCNLGSEW 31837  
QY 209 ----- 208  
Db 31838 SEISEPITPKSDVPIQAPHFKEELNLNVRYQSNATLVCKVTGHPKPIVKWYRQKEIIA 31897  
QY 209 ----- 208  
Db 31898 DGLYRIOEFKGYHOLIIASVTDDATVYQVRATNOGGSVSGTASLEVEVPKIHLPKT 31957  
QY 209 ----- 208  
Db 31958 LEGMAVHALRGEVSVKIPFSGKDPVITWQKQDLIDNNGHYQVIVTRFSTSLVFPNG 32017  
QY 209 ----- 208  
Db 32018 VERKDAGYVVCANRFGIDQKTVELDVADVPDPGRGVKVDSDVSNLVTWTEPASDGG 32077  
QY 209 ----- 208  
Db 32078 SKITNIVIEKATTAERWLRVGOARETRYTVINLFGKTSYOFRTVAENKFGLSKPSPE 32137  
QY 209 ----- 208  
Db 32138 PTITKDKTRAMVDEEVDRETVSMKASHSSTKELYEKYMAIEDLGRGEFGIVHRCVE 32197  
QY 209 ----- 212  
Db 32198 TSSKKTYMAKGVKGTDOVLVKKELISILNIAHRNLIHLHESFESMEELVMIFETISGL 32257  
QY 213 ----- 216  
Db 32258 DIFERINTSAFELNEREIVSVHOVCEALQFLSHNIGHFDIRPENIIYQTRRSSTIKII 32317  
QY 217 ----- 216  
Db 32318 EFGAOLKPGDNFRLLFTAPEYAPEVHOHDVVSTATDMWSLGTLYVLLSGINPFLAE 32377  
QY 217 ----- 218  
Db 32378 TNQOIENIMNAEYTFDEAFKEISIEAMDFVDRLLVKERKSRMTASEALQHPWLKOKIE 32437  
QY 219 ----- 225  
----- QKSLSLS -----

Db 32438 RVSTKVIRTLKRRYYHTLIKLDLNMVSAARISCGGAIRSQKGVSAKVVKVASIEIGPV 32497  
QY 226 ----- 225  
Db 32498 SQQIMHAVGEGGHVYCKIENYDQSTQVTWYFGVROLENSEKEYEITYEDGVAILYVKD 32557  
QY 226 ----- 225  
Db 32558 ITKLLDDGTYRCKVNDYGEDSSYAEFLVKGVREYDYVYCRRTMKKIKRRTDMRLLEPP 32617  
QY 226 ----- 225  
Db 32618 EFTLPLYNKTAAYGVENVRFVGTITVHPEPHVTWYKSGOKIKPGDNDKXYTPESDKGLQL 32677  
QY 226 ----- 225  
Db 32678 TINSVTTDDDAEYTVVARNKYGEDSKAKLTVTLHPPPTDSTLRPMFKRLLANABCEQ 32737  
QY 226 ----- 225  
Db 32738 SVCFEIRVSGIPPTLKWKGQPLSLGPNIEI IHGLDYVALHTRDLPEDTGYRYRTA 32797  
QY 226 ----- 225  
Db 32798 TINTAGSTSCQAHLOVERLYKKQEFKKEEHERHVQKQIDKTLRMAEILSGTESVPLTQV 32857  
QY 226 ----- 225  
Db 32858 AKEALREAAVLYKPAVSTKTVKGEFRLEIEBKKEERLMPYDVPEPRKYKQTTIEEDQR 32917  
QY 226 ----- 228  
Db 32918 IKQFVPMDSMKYKKIRDOYEMPGKLRVOKRPRIRLSRWEQFYVMPPLPRITDQYRPK 32977  
QY 229 ----- 228  
Db 32978 WRIPKLSQDDLEIVRPARRRTPSPDYFYRPRRRSLGDISDEELLPLIDDYLANMKRTEE 33037  
QY 229 ----- 228  
Db 33038 ERLLEELELGFASPPSRSPPHFELSSRLSYSPQAHVKVETRKDFRSTYHIPTKAE 33097  
QY 229 ----- 228  
Db 33098 ASTSYAELRERHAQAAVROPKORIMAREDEDELLRPVTTTQHLSEYKSELDPMKSEK 33157  
QY 229 ----- 228  
Db 33158 SRKKSRRQREYTEITEIEEYEISKHAQRESSSSASRLRRRRSILSPITYIELMRPVSELI 33217  
QY 229 ----- 228  
Db 33218 RSRPQAEYEDDTERRSPTERTPRSPSPVSSERSLSPERSAREDFISRYESMKAAL 33277  
QY 229 ----- 228  
Db 33278 KTQKTSERKEYVLSQOPFTLDHAPRITLRMRSHRVPQGNTRFILNVOSKPTAEVKWYHN 33337  
QY 229 ----- 228  
Db 33338 GVELOESSKIHTNTSGVLTLEILDCHTDGSGTYRACVNTYKGEASDYATLDVTGGDYTT 33397  
QY 229 ----- 228  
Db 33398 YASORDEEVPVFPPELRTREAYAVSFKTSEMEASSVREVYKQMTETRESLSYEH 33457  
QY 229 ----- 228  
Db 33458 SASAEMKSAALEKSLSEKSTTRKIKITTLAARILTKPRSMTVYEGESARFSCDGDGPVP 33517  
QY 229 --DWL----- 231  
-----



Db 29138 RFKVTGLTEGNEYEFHVMAENAGYGPASGISRLIKREPVPNPGPPTWVKVTDTSKTTV 29197  
QY 183 ----- 182  
Db 29198 SLEWSKPVFDGMEIIGYIIEMCKRADLGDWHKVNACVKTRYTVTDLOAGEEYKFRVSA 29257  
QY 183 ----- 182  
Db 29258 INGAGKGDCEVTGTIKAVDRLTAPELDIDANFKQTHVVRAGASIRLFIAOGRPTTAV 29317  
QY 183 ----- 182  
Db 29318 WSKPDSNLSRADIHHTDSFSTLTVENCNRNDAGKYTLTVENNSGSKSITFTVKVLDTGP 29377  
QY 183 ----- 182  
Db 29378 PPGPITFKDVTGTSATLMDAPLLDGGARHHYVVEKREASRSWQVISEKTRQIIFKVN 29437  
QY 183 ----- 182  
Db 29438 DLAEVGPYIFRVSANVEYGVGEPEYEMPEPIVATEQPAPPRRLDVVDTSKSAVLAWLKP 29497  
QY 183 ----- 182  
Db 29498 HDGSRITGYLLEMQRKGSDFWVEAGHTKOLTFTEVERLVEKTEYEFVRKAKNDAGYSEPR 29557  
QY 183 ----- 182  
Db 29558 EAFSSVILKEPQIEPTADLTGINTLITCKAGSFTTIDVPISGRPAPKVTKWLEMLKE 29617  
QY 183 -----GSFFL----- 187  
Db 29618 TDRVSITTKDRTTLTVKDSMRGDSGRYFLTLENTAGVKTFSTVVVIGRPGVTPGIEV 29677  
QY 188 ----- 187  
Db 29678 SSVSAESCVLSWGEKPDGCGTEITNYIVEKRESGTTAWOLVNSVKRTQIKVTHLTKYME 29737  
QY 188 ----- 187  
Db 29738 YSRVSENRFVSKPLESAPIIAEHFPVPPSAPTRPEVHVHSANAMWSIRWEEFYHDGGS 29797  
QY 188 ----- 187  
Db 29798 KIIGYWEKKERNTILWVKENKVPCLCNKYVTGLVEGLEQFRTYALNAGYSKASEAS 29857  
QY 188 ----- 187  
Db 29858 RPIAQNPDVAPGRPEVTDVTRSTVSLWSAPAYDGGKVVGYYIERKPVSEVGDGRWLK 29917  
QY 188 ----- 187  
Db 29918 CNYTIYSDNFFVTALSEGDTEFRVLAKNAGVISKSESTGPVTCRDEYAPPKAELDA 29977  
QY 188 ----- 187  
Db 29978 RLHGLVTRAGSDLVDAAGVKPEPKIITWKGDELCEKVSLOYTGKRATAVIKFC 30037  
QY 188 ----- 187  
Db 30038 DRSDSGKYTLTVKNASGTRKAVSMVKVLDSPGCGKLTVSRTVQEKCTLAWSLPQEDGGA 30097  
QY 188 ----- 187  
Db 30098 EITHYIVERRETSRLNWIVEGECPTLSYVVTRLIKNNYIFRRAVNKYGPGVPESEP 30157  
QY 188 ----- 187  
Db 30158 IVARNSFTIPSPGPIPEEVGTGKEHIIQWTKPESDGGNEISNYLVDKREKSLRWTRVN 30217  
QY 188 ----- 187  
Db 30218 KDYVVYDTRLKVTSLMEGCDYQFRTAVNAAGNSBPSEASNFSICREPSYTPGPPSAPRV 30277

QY 188 ----- 187  
Db 30278 VDTTKHSISLAWTKPMYDGGTDIVGYVLEMOEKDQDQWVRVHTNATIRNTEFTVPLDKMG 30337  
QY 188 ----- 187  
Db 30338 QKYSFRVAANVWKGMSSEYSAIEIEPVERIEIPDELAADDLKKTVTIRAGASLRLMSV 30397  
QY 188 -----YSKLTVDK----- 195  
Db 30398 SGRPPPVITWSKQIGIDLASRAIIDTTESYLLIVDKVNRYPDAGKYTIEAENOSGKKSATV 30457  
QY 196 ----- 195  
Db 30458 LVKYDTPGPCPSVKVKEVSRDSVTITWEIPTIDGAPVNNYIVEKREAAAMRAFKVTVTK 30517  
QY 196 ----- 195  
Db 30518 CSKTLYRISGLVEGTMYFRVLPENIYIGIBPCETSDAULVSEVPLVPAKLEVVDVTKST 30577  
QY 196 ----- 195  
Db 30578 VTLAWEKPLYDGGSRLAGYVLEACKAGTERMMKVVTLLKPTVLEHTVTSLNEGEQYLFPR 30637  
QY 196 ----- 195  
Db 30638 AQNEKGVSEPRETVTAVTVQDLRVLPTIDLSTMPQKTIHVPAGRPVELVPIAGRPPAA 30697  
QY 196 ----- 195  
Db 30698 SWFFAGSKLRESERVTVETHTKVAKLTIRETTIRDTGEYTLKKNVTGTTSETIKVILD 30757  
QY 196 ----- 195  
Db 30758 KPGPTGPIKIDEIDATSIISWPEPDLGAPLSGVVFEQDARHRCWLPVSESVRST 30817  
QY 196 ----- 195  
Db 30818 FKFTLREGNEYFRVAATNRFVIGISYLOSEVIECRSSIRIPGPETLQIFDVSRDGMTL 30877  
QY 196 ----- 195  
Db 30878 TWTPEDDGGSQVTGYIVERKEVRADRWVRNVKVPVTMTRYRSTGLTEGLEYHRVTAIN 30937  
QY 196 ----- 195  
Db 30938 ARGSGKPSRPSKPIVAMDPIAPPKQNPVTDTRTSVSLAWSVPEDEGSKVTGYLIE 30997  
QY 196 ----- 195  
Db 30998 MOKVDQHEWTKCNTTPTKIREYTLTHLPQGAERYFRVLACNAGGPGPEAEVPGTVKVTM 31057  
QY 196 -----SRWQOG----- 201  
Db 31058 LEYDPYELDERYQEGIFVRQGGVIRLTIPIKGPFPICKWTKEGODISKRAMIATSETHT 31117  
QY 202 ----- 201  
Db 31118 ELVIKEADRGDSYDLVLNCKCGKAVYKVRVIGSPNSPEGPLEYDDIQVRSVRVSWR 31177  
QY 202 ----- 201  
Db 31178 PPADGGADILGYLERREVVPKAAWYTIIDSRVGTSLVVKGLKENYEHFRVSAENQFGI 31237  
QY 202 ----- 201  
Db 31238 SKPLKSEEPVTPKTLPLNPPEPSPNPPEVLDVTKSSVLSWSRPKDDGSRVTGYIERKE 31297  
QY 202 ----- 201  
Db 31298 TSTDKVRHNKTOITTTMTVTGLVDPDAEYQFRIIAQNDVGLSETSPASEPVVCKDPFDK 31357

QY 176 ----- 175  
Db 26978 CEYEVVAENAGLSLPSETSPILRAEDPVLPSPPKIVDSGKTTITIAWKPLFD 27037  
QY 176 ----- 175  
Db 27038 GGAPITGYTVEYKSDTDWKTISIOLRGTEVTISGLTTGAETVFRVKSVMKVGASDPDS 27097  
QY 176 ----- 175  
Db 27098 SSDPOIAKEREEPLFDIDSEMRKTLIVKAGASFTMTVPFRGRPVPNVLSKPPDLDLKR 27157  
QY 176 ----- 175  
Db 27158 AYVDTTDSRTSLTIENANRDSGKYTLTIQNVLSAASLTLVVKVLDTPGPTNIIVQDVT 27217  
QY 176 ----- 175  
Db 27218 KESAVLSWDVPENDGGAPVKNYHIEKREASKAMVSVTNCNRLSYKVTNLQEGAIYYFR 27277  
QY 176 ----- 175  
Db 27278 VSGNEFGVIPAETKEGVKITEKPSPEKLGVTISKDSVSLTWLKPEDGGSRIHVY 27337  
QY 176 ----- 175  
Db 27338 VEALEKGOKNWKCAVAKSTHHVYSGLRENSSEYFRVFAENQAGLSDPRELLLPVLKEQ 27397  
QY 176 ----- 175  
Db 27398 LEPPEIDMKNFPHSTVYVRAGSNLKVDPISGKPLPKVTLSDRGVPLKATMRNFTEITAE 27457  
QY 176 ----- 175  
Db 27458 NLTLNKESVTADAGRYEITAANSSGTTKAFINIVLDRPGPTGPVVISDITEESVTLK 27517  
QY 176 ----- 175  
Db 27518 WEPPYDGSQVTNYILLKRETSTAVTEVSATVARTVMKYNKLTGEEYQFRIKAENRF 27577  
QY 176 ----- 175  
Db 27578 GISDHDSACVTVKLPYTPGPPSTPWVNTVRESITVGWHEPVSNSSAVGVYHLEKD 27637  
QY 176 ----- 175  
Db 27638 RNSILWQANKLVIRTHFKVTTISAGLIYEFVRVAENAGVGKPSHPSEPVLAIDACEP 27697  
QY 176 ----- 175  
Db 27698 PRNVRTDISKNSVLSWQQPAFDGSKIITGYIVERRDLPDGRWTKASFTNVTTQFIIS 27757  
QY 176 ----- 175  
Db 27758 GLTONSQYEFVRFARNAGVSGISNPSEVGPITCIDSYGGPVIDLPLEYEVVKYRAGTSV 27817  
QY 176 ----- 175  
Db 27818 KLRAGISGKPAPTIEWYKDDKELOTNALVCVENTDILASILIKDADRNSGCEYELKURNA 27877  
QY 176 ----- 175  
Db 27878 MGSATIRVQILDKPGPPGPIEFKVTAEKITLLWRPPADDGAKITHYIVEKRETSR 27937  
QY 176 ----- 175  
Db 27938 VVMSVSEHLEECIITTTKIIGNEYIFRVRANKYIGEPLESDSVVAKNAFVTPGPPG 27997  
QY 176 ----- 175  
Db 27998 IPEVTKITKNSMTVVWSRPTADGSDISGYFLEKRDKSLGWFKVLKETIRDTROKVTGL 28057  
QY 176 ----- 175

Db 28058 TENSQYQYRVCANAGOGPFSEPFYKAADPIDPPGPAKIRIADSTKSSITLWGSKP 28117  
QY 176 ----- 175  
Db 28118 VYDGSVAVTGYVVEIROGEEEMTTVSTKGEVRTTEYVVSNLKPGVNYFRYSAVNCAGQ 28177  
QY 176 ----- 175  
Db 28178 GEPIEMNEPVQAKDILEAPEIDLDVALRTSVIAKAGEDVQVLI PFKGRPPPTVTRKDEK 28237  
QY 176 ----- 175  
Db 28238 NLGDARYSIENTDSSLLTIPOVTRNDTGKYLITENGCGEPKSTVSVKVLDTPAACQ 28297  
QY 176 ----- 175  
Db 28298 KLQVKHVRSGVTLLWDPPLIDGSPIIINYVIEKRDATKRTWSVVSHKCSSTSFKLIDL 28357  
QY 176 ----- 175  
Db 28358 EKTPEFFRVLAEINEIGIGECETTEPVKAAEVPAPIRLSMKDSTKTSVLSWTKPDFG 28417  
QY 176 ----- 175  
Db 28418 GSVITEYVVERKGEGEQTWSHAGISKTCIEIVSQLKEQSVLEFRVFAKNEKGLSDPVTIG 28477  
QY 176 ----- 175  
Db 28478 PITVKELIITPEVDLSDIPGAQVTVRIGHNVHLELPYKGPSPISMLKDGLPKSEFV 28537  
QY 176 ----- 175  
Db 28538 RFSKTENKITLSIKNAKKEHGKGYTVILDNAVCRIVAPITVITLPPSPKPGKPIRDEIK 28597  
QY 176 ----- 175  
Db 28598 ADSVILSWDVPDNGGGEITCYSIEKRETSQTNMKWCVSSVARTTFKVPNLVKDAEYQFR 28657  
QY 176 ----- 175  
Db 28658 VRAENRYGVSQPLVSSIIIVAKHQFRIPGPGKPVIIYNTSDGMSLTWDAPVYDGGSEVTG 28717  
QY 176 ----- 175  
Db 28718 FHVEKKERNILWOKVNTSPISGREYRATGLVEGLDYQFRVYAENSAGLSSPSPDPSKFTL 28777  
QY 176 ----- 175  
Db 28778 AVSPVDPGTPDYIDVTRETITLKNPPLRDGGSKIVGYSIEKROGNERWVRCNFTDYSE 28837  
QY 176 ----- 178  
Db 28838 CQYTVTGLSPGDRYEFRIIARNAVGTISPPSQSGIIMTRDENVPPIVEFGPEYFDGLII 28897  
QY 179 ----- 178  
Db 28898 KSGESLRIKALVQGRPVPRVTFWKGVEIEKRMNMEITDVLGSTSLFVRDATRDRHGVYT 28957  
QY 179 ----- 178  
Db 28958 VEAKNASGAKAEIKVKVQDTPGKVVGPIRFTNITGEKMTLWMDAPLNDGCAPITHYIIE 29017  
QY 179 ----- 182  
Db 29018 KRETSRLAWALIEDKCEAQSYTAIKLINGNEYQFRVSAVNFVGVGRPLDSDPVVAIQYT 29077  
QY 183 ----- 182  
Db 29078 VPDAPGIPSPNITGNSITLTWARPESDGGSEIOQYILERREKKSTRWVKVSKRPISET 29137  
QY 183 ----- 182

Db 24758 AKDSMTVCNRPDSGGSEIIIGYIIVEKRDGIRWIKCNKRITDLRLRVTLGTEDHEYE 24817  
QY 168 ----- 167  
Db 24818 FRVSAENAGVGEPSATVYKACDPVKPGPTNAHIVDTTKNSITLAWGKPIYDGGSE 24877  
QY 168 ----- 167  
Db 24878 ILGYVVEICKADEEEMQIVTQOTGLRVTRFEISKLEHGEYKIRVCALNKVGLGEATVP 24937  
QY 168 ----- 167  
Db 24938 GTVKPEDKLEAPELDSELKRGIVVRAGGSARIHIPKGRPTPEITWSREGEFTDKVQ 24997  
QY 168 -----PEN----- 170  
Db 24998 IEKGVNYTOLSIDNCRNDAGKYILKLENSGSKSAFVTVKVLDTGPPQNLAVKEVRKD 25057  
QY 171 ----- 170  
Db 25058 SAFLWEPPIIDGAKVKNYVIDKRESTRKAYANVSSKSTSKFVENLGEAIYFRVM 25117  
QY 171 ----- 170  
Db 25118 AENERGVGPVETVDAVKAAPPPPGKVTLDVQSASLWKEKPEHDGGSRLGYVVE 25177  
QY 171 ----- 170  
Db 25178 MQPKGTEKWSIVAESKVCNAVTVGLSSGOEYQFRVKAYNEKSKSDPRVLGVPVIAKDLTI 25237  
QY 171 ----- 170  
Db 25238 QPSLKLPFNTYSIQAGEDLKIEIPVIGRPRPNISWVKDGEPLKQTRVNVEETATSTVLH 25297  
QY 171 ----- 170  
Db 25298 IKEGNKDDFGKYVTATNSAGTATENLSVIVLEKPGPPVGVPRFDEVSADFVVISWEPPA 25357  
QY 171 ----- 170  
Db 25358 YRGQCISNYIIVEKRDTTTTTHWMSATVARTTIKITLKTGTEYQFRIFAENRYGKSAP 25417  
QY 171 ----- 170  
Db 25418 LDSKAVIOYPKPEPGPTPEVTISIKDOMLVQWHEPVNDGGTKIIGVHLEQKEKNSIL 25477  
QY 171 ----- 170  
Db 25478 WVKLNKTPIDTKFKTTGLDEGLEEYEFKVSANIVIGIKPSKVSCEFVARDCDPPGRPE 25537  
QY 171 ----- 170  
Db 25538 AIVITRNNVTWKWKPAYDGGSKITGYIIVEKKDLDPGRMKASFNVLETFVTVSLVED 25597  
QY 171 ----- 170  
Db 25598 QRYEFVRIARNAAGNFSEPSDSSGAITARDEIDAPNASLDPKYKDVIVVHAGETVLEAD 25657  
QY 171 ----- 170  
Db 25658 IRGKPIPDVWWSKSGKELEETAARMEIKSTIQKTTLVVKDCIRTDGGQVILKLSNVGGTK 25717  
QY 171 ----- 170  
Db 25718 SIPITVKVLDPRGPPGPKVGTVAEKYLANWPPLODGGANISHYIIEKRETSRLSWT 25777  
QY 171 -----NYKT----- 175  
Db 25778 QVSTEVOALNYKVTLLPGNEYIFRVMVAVNKYIGIGEPLESGPVTACNPKPPGPPSTPEV 25837  
QY 176 ----- 175  
Db 25838 SAITKDSMVVWARVPVDDGGTEIEGYILEKKDKGVWTKCNKKTLDLRLRVTLGTBEGH 25897

QY 176 ----- 175  
Db 25898 SYEFRVAENAGVGEPSVYFIRACDALYPPGPPSNPKVTDTSRSSVSLAWSKPIYDG 25957  
QY 176 ----- 175  
Db 25958 GAPVKGYYVVEKAAADEWTTCTPTPTGLOGKQFTVTKLKENTEYFRICAINSEGVGEPA 26017  
QY 176 ----- 175  
Db 26018 TLPGSVVAQERIEPPEIELDADLRKVYVLRASATLRLFTIKGRPEPEVKEKAEGILTD 26077  
QY 176 ----- 175  
Db 26078 RAOIEVTSFTMLVIDNVTREDSGRYNLTLENNSGSKTAFVNVRLDSPSPVNLTIREV 26137  
QY 176 ----- 175  
Db 26138 KKDSVTLSEPPPLIDGGAKITNYIVEKRETTTKAYATITNNCTKTTFRIENLOEGSYIF 26197  
QY 176 ----- 175  
Db 26198 RVLASNEYCIGLPAETTEPVKYSEPPPLPGRVTVLVDVTRNTATIKWEKPESDGGSKITGY 26257  
QY 176 ----- 175  
Db 26258 VVEMQTKGSEKWTCTQVKTLEATISGLTAGEEYVFRVAANEKGRSDPRQLGVPVIARD 26317  
QY 176 ----- 175  
Db 26318 IEIKPSVELPFTFNVNKAQOLKIDVPFKGRPOATVNNRKDGOTLKETTRVNVSSKVTY 26377  
QY 176 ----- 175  
Db 26378 SLISKEASKEDVGTVELCVSNSAGSITVPTIIVLDRPQPGPIRIDEVSCDSITISWNP 26437  
QY 176 ----- 175  
Db 26438 PEYDGGCISNYIIVEKKTETTTTHWIVSOAVARTSIKIVRLTTGSEYQFRVCAENRYGKS 26497  
QY 176 ----- 175  
Db 26498 SYSSESAVVAEYFPSPGPPGPKVHVHATKSTMLVTVQVNVNDGGSRVYGYHLEYKERS 26557  
QY 176 ----- 175  
Db 26558 ILWSKANKILIADTQMKVSGLDEGLMYEYRVAENIAGIGCKSCPEVPARDPCDPPGQ 26617  
QY 176 ----- 175  
Db 26618 PEVTNITRKSLSKWSKPHYDGGAKITGYIVERRELDPGRWLKCNVTNIQETVFEVTELT 26677  
QY 176 ----- 175  
Db 26678 EDQRYEFVRFARNAADSVSESESTGPIIVKDDVEPPRVMDVKFRDVIIVVRAGEVLKIN 26737  
QY 176 ----- 175  
Db 26738 ADIAGRPLPVISWAKDGEIEERARTEIISTDNHLLTVKDCIRRDYGOVVLTKNVAGT 26797  
QY 176 ----- 175  
Db 26798 RSVAVNCKVLDKPPGPPAGPLEINGLTAECSLSGWRQBDGGADIDYIYIEKRETSRLAW 26857  
QY 176 ----- 175  
Db 26858 TICGELOMTSCKVTLLKAGNEYIFRVTGVNKYGVGEPLSAIKALDPFTVPSPPTSLE 26917  
QY 176 ----- 175  
Db 26918 ITSVTKESMTLWNSRPESDGGSEISGYIIERREKNSLRVRVNNKPVYDLRVKSTGLREG 26977

QY	168	-----	167
Db	22598	MVVCWGHDPDSGGSEIINIVERRDKAGQRIKCNKKTLDLRYKVSGLTEGHEYEFRI	22657
QY	168	-----	167
Db	22658	AENAAIGSAPSPSPFYKACDVFVKPGPGNPRVLDTSRSSISIAMNKPIYDGGSEITGY	22717
QY	168	-----	167
Db	22718	MVEIALPEDEWQIIVTPAGLKATSYITIGLTENQEKIRIYAMNSEGLGEPALVPGTPK	22777
QY	168	-----	167
Db	22778	AEDRMLPETELDADLRKVVITIRACCTLRFPVPIKGRPAPEVKWARDHGESLDKASTEST	22837
QY	168	-----	167
Db	22838	SSYTLIVGNVRPDSGKIILTVENSSGSKSAFVNRVLDTGPPQDLKVKVTKTSVTL	22897
QY	168	-----	167
Db	22898	TWDPPLLDGGSKIKNYIVEKRESTRKAYSTVATNCHTSMKWDLQLEGCSYYFRVLAENE	22957
QY	168	-----	167
Db	22958	YGIGLPAETAESVKASERPLPGKILMDVTRNSVLSWEKPEHDGSRILGIYIVEMQTK	23017
QY	168	-----	167
Db	23018	GSDKWATCATVKVTEATITGLIOGEYSFRVSAQNEKGISDPRQLSVPIAKDLVIPPAP	23077
QY	168	-----	167
Db	23078	KLLFNTFTVLAGEDLKVDVDFIGRPPTPAVTWHKDNVPLKOTTRVNAESTENNSLLTIKDA	23137
QY	168	-----	167
Db	23138	CREVDGVHVVKLTNSAGEAIETLNVILDKPGPTGPVKMDEVTAADSIILSWGPPKYDGG	23197
QY	168	-----	167
Db	23198	SSINNYIVEKRDRTTWTQIVSATVARTTIKACRLKTGCEYQFRIAAENRYGKSTYLNSE	23257
QY	168	-----	167
Db	23258	PTVAQYPFKVPGPPTPVVTLSSRDSMEVQWNEPISDGGSRVIGYHLERKERNISILWVKL	23317
QY	168	-----	167
Db	23318	NKTIPIQTKFTTGLEBGEVEFRVSAENIVGIGKPSKVSCEYVARDPCDPPGRPEAIIV	23377
QY	168	-----	167
Db	23378	TRNSVTLOWKKPYDGGSKITGYIVEKKELPEGRWMKASFTNIIDTHFEVTLGVEDHRYE	23437
QY	168	-----	167
Db	23438	FRVIARNAAGVFSESESTGAITARDEVDPPIRISMDPKYKDTIVVHAGESFKVDADIYK	23497
QY	168	-----	167
Db	23498	PIPTIOWIKDQELSNARLEIKSTDFATSLSVKDAVRVDSGNYILKAKNVAGERSVTYN	23557
QY	168	-----	167
Db	23558	VKVLDRPGPEGPWISGVTAERKCTLAWKPPLODGGSDIINIYIVERRETSLRWTVVDAN	23617
QY	168	-----	167
Db	23618	VQTLSCVKTKLEGEYTFRIMAVNKYGVGEPLESEBPVAKNPFVVPDAPKAPVETVTVK	23677
QY	168	-----	167

Db	23678	DSMIVVWERPASDGGSEILGVLEKRDKEGIRWTRCHKRLIGELRLRVTGLIENHDEYFR	23737
QY	168	-----	167
Db	23738	VSAENAGLSEPPSPSAIQKACDPIYKPGPNPNPKVIDITRSSVFLSWSKPIYDGGCEIQ	23797
QY	168	-----	167
Db	23798	GYIVEKCDVSGEWTMCTPPTGINKTNIEVEKLLLEKHEYNFRICAINKAGVGEHADVPGP	23857
QY	168	-----	167
Db	23858	IIVEEKLEAPDIDLLELRKIIINIRAGGSLRFLVPIKGRPTPEVKMGKVGDEIRDAAIID	23917
QY	168	-----	167
Db	23918	VTSFTSLVDNVRVDSGKYTITLSENSGTSKSAFVTVRVLDTPSPVNLKVTEITKDSV	23977
QY	168	-----	167
Db	23978	SITWEPPLDGGSKIKNYIVEKREATRKSAAVVTNCHKNKSWKIDLOEGCSYYFRVTAE	24037
QY	168	-----	167
Db	24038	NEYGIGLPAQTADPIKVAEVPQPGKITVDVTRNSVLSWTKPEHDGSGKIIQYIVEMQ	24097
QY	168	-----	167
Db	24098	AKHSEKSECARVKSQAIVITNLQTGEEYLFRRVVAVNEKGRSDPRSLAVPIVAKDLVIEP	24157
QY	168	-----	167
Db	24158	DVRPAFSSYVQGDULKIEVPISSGRPKPITWTKQGLPLKQTRINVTDSLOLTLUSIK	24217
QY	168	-----	167
Db	24218	ETHKDDGGQYITVANVVGQKTASIEIVTLDKPDPKPGVKFDDVSAESITLSWNPPLYT	24277
QY	168	-----	167
Db	24278	GGQITNYIVQKRDTTWTWVDVVSATVARTTLKVTCLKTGTETQFRIFAENRYGQSPALE	24337
QY	168	-----	167
Db	24338	SDPIVAQYPIKEPGPGTTPFATAISKDSWVIQWHEPVNNGGSPVIGYHLERKERNISILWT	24397
QY	168	-----	167
Db	24398	KVNKTIIDHTQFKAQNLLEGIEYEFVYAENIVGVGKASKNSECYVARDPCDPPGTPEPI	24457
QY	168	-----	167
Db	24458	MVKRNEITLOWTKPVYDGGSMITGYIVEKRDLPDGRWMKASFTNVIETQFTVSGLTEDOR	24517
QY	168	-----	167
Db	24518	YEFRTAKNAAGAIKSPSDSTGPITAKDEVLPRIKSMDFRDTIIVNAGETFRLEADVH	24577
QY	168	-----	167
Db	24578	GRPLPTIWLGRGDKIEESARCEIKNTDFKALLIVKDAIRIDGGQYILLRASNVAGSKSFP	24637
QY	168	-----	167
Db	24638	VNVKVLDRPGPEGPVQVGTGTSEKSLTWSPPLODGGSDISHYVVEKRETSRLAWTVVA	24697
QY	168	-----	167
Db	24698	SEVVTNSLKVTKLEGEYTFRIMAVNKYGVGEPLESAPVLMKNPFLVPGPPKLSLEVTNI	24757
QY	168	-----	167

Db 20378 KLIEGHEYQFRICAENKYGVGDPVTEPAIAKNPYDPPGRCDPPVISNITKDHMTVSWKP 20437  
QY 154 ----- 153  
Db 20438 PADDGSPITGYLLEKRETOAVNWKVNKKPIIERTLKATGLOEGTEYEFRTAINKAGP 20497  
QY 154 ----- 153  
Db 20498 GKPSDASKAAAYARDQYPPAPPAFPKVDYDTRSSVSLSWGKPAYDGGSPILGYLVEVKRA 20557  
QY 154 -----YPSDIAY- 160  
Db 20558 DSDNNVRCNLPONLOKTRFEVTLGMDTQYQFRVAVNKNIGYSDPSDVPDKRHYPKDILIP 20617  
QY 161 -----|| || : 160  
Db 20618 PEGELDADLRKTLILRAGVTWRLYVPVKGRPPKITWSKPNVNLDRICLDIKSDFTDFF 20677  
QY 161 ----- 160  
Db 20678 LRCENVKYDAGYILTLSENSCGKREYTIIVVKVLDTPGPPVNVTVKEISKDSAYVTWEP 20737  
QY 161 ----- 160  
Db 20738 IIDGGSPINIYVQRDAERKSNSTVTTECKSTSPRVANLEBEGKSYFFRVFAENYIGID 20797  
QY 161 ----- 160  
Db 20798 PGETRAVAKASQTPGVVVDLKVRSYSKSSCSIGWKKPHSDGSGRIIGYVVDLFTENKWQ 20857  
QY 161 ----- 160  
Db 20858 RYMKSLQYSAKOLTEGKEYTFRVSAENENGEPSEITVVARDVAVAPDLKGLPDL 20917  
QY 161 ----- 160  
Db 20918 CYLAKENSFRLLKIPKGPAPSVWKKGEDPLATDTRVSVESAVNTTLIVDCOKSDA 20977  
QY 161 ----- 160  
Db 20978 GRYTITLNVAGTKEGTISIKVVGKPGIPTGPIKFEVTAEAMTLKWAPPKDDGSEITN 21037  
QY 161 ----- 160  
Db 21038 YILEKRDSVNNKWTCASAVQKTTFRVRLHEGMEYTFRVAENKYGVGEGLKSEPIVAR 21097  
QY 161 ----- 160  
Db 21098 HPFDVPDAPPPNIYDVRHDSVSLTWDPKKTGGSPITGYHLEFKERNLLWKRAKTP 21157  
QY 161 ----- 160  
Db 21158 RMRDFKVTGLTEGLEYEFVRMAINLAGVKPSLPSEPVALDPIDPPGKPEVINITRNSV 21217  
QY 161 ----- 160  
Db 21218 TLIWTEPKYDGHKLTGYIVEKRDLPKSMKANHVNVPECAFTVTDLVEGKYEFIR 21277  
QY 161 ----- 160  
Db 21278 KNTAGIASPSESTETIICKDEYEAPTIVLDPTIKDGLTIKAGDTIVLNAISILKPLPK 21337  
QY 161 ----- 160  
Db 21338 SWSKAGDIRPSDITQITSTPSSMLTIKYATRKDAGEYITATNPFGTKVEHVKVTVL 21397  
QY 161 ----- 160  
Db 21398 DVPGGPPEVSNVSAEKATLTWTPLEDGGSPIKSYILEKRETSRLLTWTVVSEDIQSCR 21457  
QY 161 ----- 160  
Db 21458 HVATKLIQNEYIFRVSAVNHYGKEPVQSEPVKAVDRFGPPGPEKPEVSNVTNTATV 21517

QY 161 ----- 160  
Db 21518 SWKRPVDDGGSEITCYHYVERREKKSILRWVIAIKTPVSDLRCKVTGLQEGSTYEFVRSAEN 21577  
QY 161 ----- 160  
Db 21578 RAGIGPSEASDVLKDAAYPPGPPSNPHVDTTKKSASLAWGRPHYDGGLEITGYVVE 21637  
QY 161 ----- 160  
Db 21638 HOKVGEAWIKDITCTALTALRITQFVVPDLQTKENYFRISAINDAGVGEPVAPIDPVEIVER 21697  
QY 161 ----- 160  
Db 21698 EMAPDFELDAELRRTLIVVRAGLSIRIFVPKGRPAPEVTWTNDKINLKNRANIENTESPT 21757  
QY 161 ----- 160  
Db 21758 LLIIPECNRYDTGKFVMTIENPAGKSGFVNVRLDTPGPVLNLRPTDITKDSVTLHMDL 21817  
QY 161 ----- 160  
Db 21818 PLIDGSRITNYIVEKREATRKSYSYATTKCHKCTYKVTGLSEGCEYFPRVMAENEYIG 21877  
QY 161 ----- 160  
Db 21878 EPTETEPVKASEAPSPDPSLIMDITKSTVSLAWPKPKHDGSKITGYVIEAQRKGSQ 21937  
QY 161 ----- 160  
Db 21938 WTHITTVKLECVVRNLTEGEYTFQVMAVNSAGRSAPRESRPVIVKEQTMPLPELDLRI 21997  
QY 161 ----- 160  
Db 21998 YQKLVIAKAGDNKVEIPVLGRPKPTVTKKGDQILKQTVRYNFETTATSTILNINECVR 22057  
QY 161 ----- 160  
Db 22058 SDGSPYPLTARNIVGEGVDITIQVHDIPGPPTGPIKFEVSSDFVTFSWDPPENDGGVP 22117  
QY 161 ----- 160  
Db 22118 ISNYVEMRQDSTTWVELATTVIIRTYKATRLTTGLEYQFRVKAQNRVGVGPGITSACI 22177  
QY 161 ----- 160  
Db 22178 VANYFPKVPGPPTQVITAVTKDSMTISWHEPLSDGGSPILGYHYVERKERNGLMQTVSK 22237  
QY 161 ----- 160  
Db 22238 ALVPGNIFKSGSLTDGIAYEFVRVIAENMAGSKSPKSEPMALADPIDPPGKPVPLNIR 22297  
QY 161 ----- 160  
Db 22298 HTVTLKAKPEYTGFKITSYIVERKDLPNGRWLKANFSNILENEFTVSGLTEDAAEYPR 22357  
QY 161 ----- 160  
Db 22358 VIAKNAAGAI SPPSESDAITCRDDVEAPKIKVDVKFDTVILKAGEAFRLEADYSGRPP 22417  
QY 161 -----EWESNGQ- 167  
Db 22418 PTMEWSKDGKELEGTAKLEIKIADFTNLVKNKDRDSDGAYTLTATNPGFAKHIFNVK 22477  
QY 168 ----- 167  
Db 22478 VLDRGPPPEGPLAVTEVTSEKCVLSWFPPLDDGGAKIDHIYVQKRETSRLATNVTNASEVQ 22537  
QY 168 ----- 167  
Db 22538 VTKLKVTKLLKGNXYIFRVMVAVNKYCVGEPELESEPVAVNPYCPDPKPNPEVTITKDS 22597

QY	154	-----	153
Db	18218	VENPAGSKTVSVKVLVLDKPGPRDLLEVSEIRKDCSYLTWKPELDDGGSVITNYVVERRD	18277
QY	154	-----	153
Db	18278	VASAQMSPLSATSKKKSHFAKHLNEGQYLFRVAANQYGRGPFVETPKPIKALDPLHPP	18337
QY	154	-----	153
Db	18338	GPPKDLHHVDVDTESVSLVWVKPDRDGGSPITGYLVEYQEBQTDWIKFKVTNLCVYT	18397
QY	154	-----	153
Db	18398	GLOQKTYRFRVAENIVGLGLPDTTIPTECQEKLVPPSVELDVKLIEGLVVKAGTTVRF	18457
QY	154	-----	153
Db	18458	PAIIRGVVPVPTAKWTTDGSSEIKTDEHYTVETDNFSSVLTIKNCRRDTGEYQITVSNAAG	18517
QY	154	-----	153
Db	18518	SKTVAVHLTVLDVPGPPTGPIINILDVTPHEMTISWQPPKDDGSPVINIIVEKQDTRKDT	18577
QY	154	-----	153
Db	18578	WGVSSGSSKTKLKIPLHQKCEYVFRVAENKIGVGPPLDSTPTVAKHKFSPSPGKP	18637
QY	154	-----	153
Db	18638	VVTDITENAAVSWTLPLKSDGSPITGYMERREVTGKVVVNKTPIADLKFRVTGLYEG	18697
QY	154	-----	153
Db	18698	NTYEFVRFAENLAGLSPSSDPIKACRPKPPGPPINPKLKDKSRETADLWTKPLSD	18757
QY	154	-----	153
Db	18758	GGPILGYVVECQKPGTAQWRINKDELIRQCAFRVPLIEGNEYFRIKAANIVGEGEP	18817
QY	154	-----	153
Db	18818	RELAESVIAKDILHPPPEVELDVTCDVITVRVGQTIILARVKRPEPDITWTKEGKVLV	18877
QY	154	-----	153
Db	18878	REKRVLIQDLPRVELQKEAVRADHGKYIISAKNSSGHAQGSAINVNLDRPGPCQNLKV	18937
QY	154	-----	153
Db	18938	TNVTKENCTISWENPLDNGSGSEITNFIVEYRKNQKGWSIVASDVTKRLIKANLANNEY	18997
QY	154	-----	153
Db	18998	YFRVCAENKVGVPPTIETKTPILAINPIDRPGEPENLHIADKGKTEVYLKRRRDPYDGS	19057
QY	154	-----	153
Db	19058	PNLSYHVERRLKGSDDWVRHKGSIKETHYMWVDRVCVENQIYEFVRQTKNEGGEWVKTE	19117
QY	154	-----	153
Db	19118	EVVVKEDLQVPVLDLKLGLVTVKAGDTIRLEAGVRGKPPPEVAWTKDKDATDLTRSPRV	19177
QY	154	-----	153
Db	19178	KIDTRADSSKFSLTAKRSDGGKYVVVTAINTAGSFVAYATVNVLDKPGPVRLKIVDVSS	19237
QY	154	-----	153
Db	19238	DRCTVCWDPDPDDGGCEIQNYILEKCTKRMWVSTYSATVLPFGTTVTRLIEGNEYIFRY	19297
QY	154	-----	153

Db	19298	RAENKIGTGPTESKPVIAKTKYDKPGRDPPEVTKVSKEMTVVMNPPEYDGGKSITGY	19357
QY	154	-----	153
Db	19358	FLEKKEKHSTRWVPVVKNSAIPERRMKVQNLDPDHEYQFRVKAENEIGIGEPSLSPRPVA	19417
QY	154	-----	153
Db	19418	KDPIEPGPPPTNPRVVDTTKHSITLWGKRPVYDGGAPIIGYVVVEMRPKIADASPDEGWR	19477
QY	154	-----	153
Db	19478	CNAQAOLVRKEFTVTSILDENQEYEFRCQNOVYIGRPAELKEAIKPKLEPEIDLDA	19537
QY	154	-----	153
Db	19538	SMRKLIVRAGCPIRLFAIVRGRPAKVTWRKVGIDNVVRKGOVLDVTMAFLVIPNSTR	19597
QY	154	-----	153
Db	19598	DDSGKSLTLVNPAGEKAVFVNVRLDTPGPVSDLKVSDVTKTSHVSHVAPPENDGGSQV	19657
QY	154	-----	153
Db	19658	THYIVEKREADRKTWSTVTPVEVKKTSFHVNTLVPGNEYIFRVTAVNEYGPGVPTDVPKPV	19717
QY	154	-----	153
Db	19718	LASDPLSEDPKPRKLEVTEMTKNSATLAWLPLLRDGGAKIDGYITSYREEEQPADRWTEY	19777
QY	154	-----	153
Db	19778	SVVKDSLVTVLKGEKKYKFRVAARNVGVSLPREAGVYEAKEQLLPPKILMPQIITI	19837
QY	154	-----	153
Db	19838	KAGKLRIEAHVYKGPHTCKWKKGDEVVTSHLAVHKADSSSILIKDVTBKDSGYIS	19897
QY	154	-----	153
Db	19898	LTAENSSGTDQKIKVVMVNDAGPPQPPFDISDIDACSLSWHIPLEDGGSNITNYIVE	19957
QY	154	-----	153
Db	19958	KCDVSRGDNVTALASVTKTSCRVGKLIPOGEYIFRVAENRFGISEPLTSPKMWQPPFG	20017
QY	154	-----	153
Db	20018	VPSEPKNARVTKVKNDCIFVAMDRPDSGGSPIIGYLIBERKERNLLWVKANDTLVRSTE	20077
QY	154	-----	153
Db	20078	YFCAGLVEGLEYSFRIYALNKAGSPSPKPTTEYVTARMPVDPGKPEVIDVTKSTVSLIW	20137
QY	154	-----	153
Db	20138	ARPKHDGGSKIIGYFVBACKLPDGKVVRCNTAPHQIQEETATGLEEKAQYQFRAIART	20197
QY	154	-----	153
Db	20198	AVNISPPSPSDPVTILAENVPRIDLSVAMKSLTVKAGTNVCLDVTVEGKPMPTVSMK	20257
QY	154	-----	153
Db	20258	KDGTLLKPAEGIKMAMORNCTLELFSVNRKDSGDYTTAENSSGSKSATIKLVLDKPG	20317
QY	154	-----	153
Db	20318	PPASVKINWYSDRAMLSWEPLEDGGSEITNYIVDKRETSRPNWAQVSATVPIITSCSVE	20377
QY	154	-----	153

Db 15998 DYKLRYSAVNAAGEGPPGETQPVVAEPQEPFAVELDVSKGQIQIMAGKTLRIPAVVTG 16057  
QY 133 ----- 132  
Db 16058 RPVPTKVTKEGELDKRWIDNVGCTKSELIKDALKRDHGRVITATNSCGSKFAAAR 16117  
QY 133 ----- 132  
Db 16118 VEVFDPGVPVLDLKPVVTRNKMCLLNWSDPEDDGSSEITGFIERKDAKMHTRQPIETE 16177  
QY 133 ----- 132  
Db 16178 RSKCDITLLEGOEYKERVIAKNKFCGPPVEIGPILAVDPLGPPTSPERLTYTERTKST 16237  
QY 133 ----- 132  
Db 16238 ITLDMKEPRSGSGPIQGYIEKRRHDKPDFERVNKRCLPTTSFLVENLDEHQMYEVRK 16297  
QY 133 ----- 132  
Db 16298 AVNEIGESEPSLPLNVYIQDDEVPPTIKRLSVRGDTIKVKAGEPVHIPADVTLGMPKPI 16357  
QY 133 ----- 132  
Db 16358 EWSKNETVIEKPTDALQITKEVSRSEAKTELSIPKAVREDKGYTYVTASNRLGSVFRNV 16417  
QY 133 ----- 132  
Db 16418 HVEVYDRPSPRNLAVTDIAESCYLTWDAPLDNGSGSEITHVIDKRDASRKAWEVYT 16477  
QY 133 ----- 132  
Db 16478 NTAVERRYGIWKLIPNGOYEFVRVAVNKYGISDECKSDRWVIQDPYRLPGPPCKPKVLAR 16537  
QY 133 ----- 132  
Db 16538 TKGSMVSWTTPPLDNGGSPITGWLEKREEGSPYRSVRAPITKVGLKGVEFNVPRLLE 16597  
QY 133 ----- PPSR 136  
Db 16598 GVKYQFRAMAINAAGIPSPDPEVAGDPIPPGPPSPCEVKDKTKSSISLGNKPPAK 16657  
QY 137 D----- 137  
Db 16658 DGGSPIKGYIVEMQEGTGWKRVNEPKLITTCWCVPNLKELRYFRVAVNEAGES 16717  
QY 138 ----- 137  
Db 16718 EPSDTTGEIPATDIOEPEFFIDIGNQDCLVCKAGSQIRIPAVIKGRPTPKSSWEPDGKA 16777  
QY 138 ----- 137  
Db 16778 KKAMKDVHDIPEDAQLETAENSSVIIPECKRSHTGKYSITAKNKGOKTANCYKVM 16837  
QY 138 ----- 137  
Db 16838 VFGPKDLKVDITRGSCLSNKMPDDGGDRIKGYIEKRTIDGKAWTKVNPDCGSTTF 16897  
QY 138 ----- 137  
Db 16898 VVPDLLSEQQYFVRVAENRFGIGPPVETIQRTTARDPIYPPDPKIKIGLITKNTVHL 16957  
QY 138 ----- 137  
Db 16958 SWKPKNDGSGSPVTHYIVECLAWDPTGKKEAWRQCNKRDEVELOFTVEDLVEGGEYFR 17017  
QY 138 ----- 137  
Db 17018 VRANNAAGVSKPSATVPGCDQRPMPPSIDLKEFMEVEEGTNVIVAKIKGVPPFTLW 17077  
QY 138 ----- 137  
Db 17078 FKAPPKKPDNKEPVLVDYTHVNLVDDTCTLVIPQSRSDTGLYITITAVNNLGTASKEMR 17137

QY 138 ----- 137  
Db 17138 LNVLRGPPVPVGIKPFESVADQMTLSWFPKDDGSKITNVYIEKREANRKTWVHSSE 17197  
QY 138 ----- 137  
Db 17198 PKECTYTIKPLLEGHEVYFRIMAQNKYIGCEPLDSEPTARNLFSVPGAPDKPTVSSVTR 17257  
QY 138 ----- 137  
Db 17258 NSWTVNWEPEYDGGSPVTGYWLEMKDITSKRWKRVNRDPIKAMTLGVSKYKVTGLIEGSD 17317  
QY 138 ----- 137  
Db 17318 YQFRVYAINAAGVGPASLPDPATARDPIAPGPPPKVTDWTKSSADLEWSPPLKDGGS 17377  
QY 138 ----- 137  
Db 17378 KVTGYIYEKKEGKEWEGKDKVEKRTKLVVTGLKEGAFYKFRVSAVNIAGIGEPGYT 17437  
QY 138 ----- 137  
Db 17438 DVIEMKDRLVSPDLQLDASVRDRIVVHAGGVIRIIAYVSGKPPPTVTWNNERTLPQEA 17497  
QY 138 ----- 137  
Db 17498 IETTAISSMWIKNCORSHQGVYSLAKNEAGERKKTIIIDVLDVPGVGTPTPLAHLNLTN 17557  
QY 138 ----- 137  
Db 17558 ESCKLTWFSPEDDGGSPITNYVIEKRESRRAWTPVTYVTRQNAFVGLIOGKAYFFRI 17617  
QY 138 ----- ELTKNOVSLT- ----- 147  
Db 17618 AAENSIGMPPFVETSEALVIREPIVPERPEDLEKVEKVTNKTVTUTWNPVKYDGGSEIIN 17677  
QY 148 ----- 147  
Db 17678 YVLESRLIGTEKFKHVTNDNLLSRKYTVKLGEGDFEYRVSANVIGQGKPSFCTKPIT 17737  
QY 148 ----- 147  
Db 17738 CKDELAPTLHLDLFRDKLITRVGEAFALTRYSGKPKVSWFKDEADVLEDDRTHIKTT 17797  
QY 148 ----- CLV- ----- KGF- ----- 153  
Db 17798 PATLALKEIKAKRSDSGKYCVVVENSTGSRKGCQVNVVDRPGPPVGPVSFDEVTKDYMV 17857  
QY 154 ----- 153  
Db 17858 ISWKPLDDGSKITNYIIIEKKEVGKDVWMPVTSASAKTTCVSKLLECKDYIFRIHAEN 17917  
QY 154 ----- 153  
Db 17918 LYGISDPLVSDSMKAKDRFVPDAPDQPIVTEVKDSALVTWKNPHDGKPTITNYILEKR 17977  
QY 154 ----- 153  
Db 17978 ETMSKRWARTKDPITHPYTKFRVPDLLECCQYEFVRSAENEIGIGDPPSPKPVFAKDPI 18037  
QY 154 ----- 153  
Db 18038 AKSPSPVNEAIDTTCNSVDLTWQPPRHGDSKILGYIYEVOKVGEWRRANHTPESCP 18097  
QY 154 ----- 153  
Db 18098 ETKYKVTGLRDGQTYKFRVLAVNAAGESDPAHVPEVLVKDRLEPELILDANMAREQHI 18157  
QY 154 ----- 153  
Db 18158 KYGDTLRLSAILIKGVPPFKVTWKKEDRDAPTAKRIDVTPVSGSKLEIRNAAHEDGGIYSLT 18217

QY 133 ----- 132  
Db 13838 PVEIIRPPQDILEAPGADVFLAELNKKDKEVQWLNRMVVOGDKHOMSEGIHRLQI 13897  
QY 133 ----- 132  
Db 13898 CDIKPRDQGEYRFIAKDKEARAKLELAAPKIKTAQDQLVVDVGKPLTMVVPYDAYPKAE 13957  
QY 133 ----- 132  
Db 13958 AEWKENEPLSTKIDTTABQTSFRILEAKKDGKGRYKIVLQNHGKAEGFINLKVIDVP 14017  
QY 133 ----- 132  
Db 14018 GPNRLEVTETFDGEVSLAWEELTDGGSKIIGYVVERRDIKRTKWLATDRAESCEFTV 14077  
QY 133 ----- 132  
Db 14078 TGLQKGGVEYLFRRSARNRVGTGEVETONPVEARSKYDVPGLNVTITDVRFGVSLT 14137  
QY 133 ----- 132  
Db 14138 WEPPEYDGAETNYVIELRDKTSIRWDTAMTVRAEDLSATVTVVGEQEYSFRVAQNR 14197  
QY 133 ----- 132  
Db 14198 IGVKPSAATPFVKVADPIERPSPVNLTSDDQTQSSVQLKWEPPPLKDGSPILGIIER 14257  
QY 133 ----- 132  
Db 14258 CEBGKDNWIRCNKMLVPELTYKVTGLEKGNKLYRVAENKAGVSDPSEILGPLTADDAF 14317  
QY 133 ----- 132  
Db 14318 VEPTMDLSAFKOGLEVIVPNPITILVPSTGYPRPTATWCFGDKVLETGDRVMKMTLSAYA 14377  
QY 133 ----- 132  
Db 14378 ELVISPERSDKGIYTLKLENRVKTTISGEIDVNVIAFPSAPKELKFGDITKDSVHLTWEP 14437  
QY 133 ----- 132  
Db 14438 PDDGGSPLTYGVVEKREVSRTWTVMDFVTDLEFTVFDLVQGEYLFKVCARNKCGPG 14497  
QY 133 ----- 132  
Db 14498 EPAYVDEPNMSTPATVPDPPENVKWRDRTANSIFLTWDPKNDGGSRIKGIYVERCPRG 14557  
QY 133 ----- 132  
Db 14558 SDKWACGEPVAETKMEVTGLEBGKWAVRVKALNRQGASKPSRPTTEEIQAVDTQEAPEI 14617  
QY 133 ----- 132  
Db 14618 FLDVKILLAGLTVKAGTKIELPATVTKPEPKITWTAKMILKODKRITIEVNPCKSTVTI 14677  
QY 133 ----- 132  
Db 14678 VDSKRSDTGTIIIEAVNVCGRATAVVEVNVLDKPGPPAAFDITDVNESCLLTWNPPRDD 14737  
QY 133 ----- 132  
Db 14738 GGSKITNYVVERRATDSEVWHKLSSTVKDTNFKATKLIPNKEYIFRVAENMYGVGEPVQ 14797  
QY 133 ----- 132  
Db 14798 ASPITAKYQDPGPPTRLEPSDITKDAVTLTWCEPDDDGSPITGYWVERLDPDPTDKW 14857  
QY 133 ----- 132  
Db 14858 RCNKMVPKDTYRVKGLTNKKKYFRVLAENLAGPKPSKSTEPILIKDIPDPWPPGKP 14917  
QY 133 ----- 132

Db 14918 TVKDVGKTSVRLNWKTRPEHGDGAKIESYVIEMLKTGTDWVRVAEGVPTTQHLLPGLMEG 14977  
QY 133 ----- 132  
Db 14978 QEYSFRVRAVNKAGESEPSDPVLCREKLYPPSPRMLVINITKNTADLAKWTVPKED 15037  
QY 133 ----- 132  
Db 15038 GGSPIITNYIVEKRDVRRKGQTVDTTVKDKCTVPTLPTGESLVFRVAENAIQSDYTE 15097  
QY 133 ----- 132  
Db 15098 IEDSVLAKDTFTTGPPIYALAVVDVTKRHVDLAKWEPKNDGGRPIQRIYVIEKKERLGRW 15157  
QY 133 ----- 132  
Db 15158 VKAGTAGPCDPCNFRVTDVIEGTEVQFQVRAENEAGVCHPSEPTIELSIEDPTSPSPPLD 15217  
QY 133 ----- 132  
Db 15218 LHVTDAGRKHIAIAWKPEKNGGSPIGYHVEMCPVCTEKKMRVNSRPIKDLKFKVEEGV 15277  
QY 133 ----- 132  
Db 15278 VPDKEYVLRVAVNAIGVSEPSISENVVAKDPDCKPTIDLETHDIIVIEGKLSIPVPF 15337  
QY 133 ----- 132  
Db 15338 RAVPPTVSWHKDGEVKASDRLTMKNDHISAHLEVPKSVRADAGIYVITLENKLSATA 15397  
QY 133 ----- 132  
Db 15398 SINVKVIGLPGPKDIAKSDITKSSCKLTWEPPEFGGTPILHYVLERREAGRTYIPVM 15457  
QY 133 ----- 132  
Db 15458 SGENKLSWTVKDLIPNGEYFFRVKAVKVGGEYIELKNPVIADPKQPPDPVDVEVHN 15517  
QY 133 ----- 132  
Db 15518 PTAEAMTITWKPLYDGGSKIMGYIIEKIAKEERWKRCNEHLVPLITYTAKGLEEGKEY 15577  
QY 133 ----- 132  
Db 15578 QFRVRAENAAGISEPSRATPPTKAVDPIDAPKVLRTSLEVKRGDEIADASIGSPYPT 15637  
QY 133 ----- 132  
Db 15638 ITWIKDENVIVPEIKKRAAPLVRRRKGEVEEPFVPLTQRLSIDNSKKGSQRLVRD 15697  
QY 133 ----- 132  
Db 15698 SURPDHGLYMIKVENDHGIKAPCTVSVLDTPGPPINFVFEDIRKTSVLCWKPEPLDDGG 15757  
QY 133 ----- 132  
Db 15758 SEIINTYLEKKDKTKPDSEWIVVTSTLRHCKYSVTKLIEGKEYLFRVRAENRFGPGPCV 15817  
QY 133 ----- 132  
Db 15818 SKPLVAKDPGPDAPDKPIVEDVTSNMLVKWNEPKDNGSPILGYWLEKREVNSTHWSR 15877  
QY 133 ----- 132  
Db 15878 VNKSLNALKANVDGLLEGLTYVFRVCAENAAAGPGKSPDPKTAHDPISPGPPPIPRV 15937  
QY 133 ----- 132  
Db 15938 TDTSSTTIELEWEPAPNGGGEIVGYVDKQLVGTNENSRCTEKMVKRQYTVYKEIREGA 15997  
QY 133 ----- 132



Db 11618 POVTVPKKPVPKAPAVAKKPELPPVKVPEVPEKVVPEKKVPLVVPKKPEAPPKVP 11677  
QY 114 ----- 113  
Db 11678 EYPKEVWPEKKVAVPKKPEVPAKVPEVKKPVLEKPAVPVPERAESPPPEVPEPEI 11737  
QY 114 -----EK 115  
||  
Db 11738 APEEETAPEEKPVVVAEEEEVPPPAVPEPKKIPEKKVPVKKPEAPPKPEPEK 11797  
QY 116 TISK----- 119  
||  
Db 11798 VIEKPKLPPPPPPAPPKEDVKEIFOLKAIPKKKVPKQVPEKVELTLPKVPGGEK 11857  
QY 120 ----- 119  
Db 11858 KVRKLLPERKPEKBEVVLKSVLRKRPEEPEKVPKLEKVKKPAVPEPPPPKPEVEE 11917  
QY 120 ----- 119  
Db 11918 VPTVTKRERKIPEPTKVPEIKPAIPAPPEPKPEAEVKTIKPPPPVEPEPTPIAAPVTV 11977  
QY 120 ----- 119  
Db 11978 PVVGKAEAKAPEAKPKGPIKGVPKTPSPIEAERKLRPGSGGKPPDEAPFTYQL 12037  
QY 120 ----- 119  
Db 12038 KAVPLKFVKEIKDIILTESEFVGSSAIFELVSPSTAITTWMKDGNSIRESKPHFIADG 12097  
QY 120 -----AKQP- 124  
||||  
Db 12098 KDKLHIIDVQLSDAGEYTCVLRNLGNKKTSTAKLVEELPVRFVKTLEEEVTVKGOPL 12157  
QY 125 ----- 124  
Db 12158 YLSCELNKEDVWRKDKIVVEKPGRIVPGVIGLMRALTINDADTDAGTIVTVVENAN 12217  
QY 125 ----- 124  
Db 12218 NLECSKVVEVIRDWLKPIRDQHVKPGTAIFACDIAKDTPNIKWFGYDEIPAEPN 12277  
QY 125 ----- 124  
Db 12278 DKTEILRDNHLYLTIKNAMPDIAEYAVEIEGKRYPAKLTILGEREVLLKPIEDTYIE 12337  
QY 125 ----- 124  
Db 12338 KESAFDAEISEADIPGQWKLKELLRPSPTCEIKABGGKRFLLTKVKVLDQAGEVLYQA 12397  
QY 125 ----- 124  
Db 12398 LNAITTAILTVEIELDFAVPLKDVTPERRQARFECVLTREANYIWSKGPDIKSSDKF 12457  
QY 125 ----- 124  
Db 12458 DIIADGKKHILWINDSQFDEGYTAEEVGKTSARLFTVGTIRLKFMSPLEDTVKEGET 12517  
QY 125 ----- 124  
Db 12518 ATFVCELSHEKMHVWFKNDAKLHTSRTVLIISSEGTHKLEMKEVTLDDISQIKAQVKEL 12577  
QY 125 ----- 124  
Db 12578 SSTAQLKVLADOPYFVTKLHDKTAVEKDEITLKCEVSKDVPVWPKFDGEEIVPSPKYSIK 12637  
QY 125 ----- 124  
Db 12638 ADGLRILIKKADLKDGVEYCDGCKTKANVTVEARLIKVEKPLYGVFVGETAHF 12697  
QY 125 ----- 124  
Db 12698 EIELSEPDVHGQWKLKGOPLTASPDCEIITDGKKHILHLNLCOLGWTGEVSFOAANAKSA 12757

QY 125 -----REPO----- 128  
|||  
Db 12758 ANLKVKEPLIFITPLSDVKVPEKDEAKPECEVSRPEKTFRMLKGTQETITGDDREFELIKD 12817  
QY 129 ----- 128  
Db 12818 GPKHSMVIKSAFAFEAKYMFADKHTSGKLIIEGIRLKFITPLKDVTAKEKESAVFTV 12877  
QY 129 ----- 128  
Db 12878 ELSHDNIRVKWKNORLHTTRSVSMQDEGKTHSITFKDLSIDDTSOIRVEAMGMSSEAK 12937  
QY 129 ----- 128  
Db 12938 LTVLEGDPYFTGKLQDYGVEKDEVILQCEISKADAPVKWFKDGRKPKNAVAKADGK 12997  
QY 129 ----- 128  
Db 12998 KRMLILKALKSDIGOYTCDCGTDKTSGKLDIEDREIKLVRPLHSVEMETETARFETI 13057  
QY 129 ----- 128  
Db 13058 SEDDIHANWKLKGEALLQTPDCEIKEEGKIHSLVHLNCRLDQDTGGVDFQOANVKSASHLR 13117  
QY 129 ----- 128  
Db 13118 VKPRVIGLLRPLKDVTVTAGETATFDCELSYEDIPVEWYLGKGLPEPSDKVVPVRSRGKVH 13177  
QY 129 ----- 128  
Db 13178 TLTLDVLEKEDAGEVOLTAQDKFTHANLFVKPEPPVEFTKPLEDQTVVEGATAVLECEVSR 13237  
QY 129 ----- 128  
Db 13238 ENAKVWFKNGTEILKSKYEIVADGRVKLVHDCPTEDIKTYTCDAKDFKTSCLNVV 13297  
QY 129 ----- 128  
Db 13298 PPHVEFLRPLTLQVREKEMARFECELSRENAKVKWFKDGAIEIKGKYDIISKGAVRIL 13357  
QY 129 ----- 128  
Db 13358 VINKCLLDEABYSECVRTARTSGMLTVLEEAFTKLANIEVSETDTIKLVCEVSKPG 13417  
QY 129 ----- 128  
Db 13418 AEVWYKGDDEIIEITGRYEILTEGRKRIIVIQNAHLEDAGNINCLRPSSRTDGKVKVHEL 13477  
QY 129 ----- 128  
Db 13478 AAEFISKPONLELEGEKAEEFVCSISKESFPVQWKRDDKTLESGDKYDVIADGKKRVLVV 13537  
QY 129 ----- 128  
Db 13538 KDATLQDMGTYYVMGAARAAHLTVIEKLRIWVPLKDRVKEQOEVEFVCEVNTGAKA 13597  
QY 129 -----VYTL----- 132  
|||  
Db 13598 KWFRNEEALFDSKYIILLOKDLVYTLIRDAHLDDQANYVSLTNHNGENVKSAANLIYE 13657  
QY 133 ----- 132  
Db 13658 EEDLRIVEPLKDIETMEKKSVTFWCKVNRNLNVLTKWNGEEVFPDNRVSYRVDKYKHML 13717  
QY 133 ----- 132  
Db 13718 TIKDCGFPDEGEYIVTAGODKSVAEILLIIEAPTEFVEHLEDQTVTEFDADFVSCOLSREK 13777  
QY 133 ----- 132  
Db 13778 ANVKWYRNGREIKGKKYKFEKDGSIHRLIHKDRLDDECEYACGVEDRKSRLRPFVEI 13837

[illegible]

Db	10538	EPKKVPPEKKVPPKVVIKMEEPLPAKVTBHRHMQITQEEKVLVAVTKEAPPKARVPBPKR	10597
Qy	114	-----	113
Db	10598	AVPESEKVLKLPRKEEDPAPAKVTGFRKRVKEEKVSIBAPKREOPIKEVTIMEEKGRAY	10657
Qy	114	-----	113
Db	10658	TLEEAHSVOREEYEEYDYKEFEVEYTEEYDQYEEYEEREYRYBEHEEYITEPE	10717
Qy	114	-----	113
Db	10718	KPIPVKPVEEPVPTPKAPPAPAKVLKKAYPEKVPVPIPKKLKPPPYPPEEKVPPEEKVFEEK	10777
Qy	114	-----	113
Db	10778	TRISITKREKQVTEPAAKVPMKPKRVAEAEKVPVRKEVAPVVRPEVKELSPBEVAF	10837
Qy	114	-----	113
Db	10838	EDEVTHVEEYLVEEBEYIHEEEBFITEEEVVPVVKVDEPKVPPEKKKVPVPKK	10897
Qy	114	-----	113
Db	10898	KEAPPAKVPPEPKPPEEKVPVLIPKKEKPPAKVPPEKVPPEKVPVPVKVEAPPA	10957
Qy	114	-----	113
Db	10958	KVPEVPKVPPEKKVPDPAPKPKVEAPPAKVEVPKKLIPPEKKKTPTVPKKVEAPPKVPK	11017
Qy	114	-----	113
Db	11018	KRPVPVPVALPOEEVLFEIEIVPEEEVLPPEEEVLPPEEEVLPPEEEIPEE	11077
Qy	114	-----	113
Db	11078	EEEEPEEYVPEEEEFVPEBEVLPEVKPKVPDPAPVPEIKKKYTKKKVJPKKEEAPPA	11137
Qy	114	-----	113
Db	11138	KVPEVPKKKEKRIILPKEEEVLPEVTEEPPEBPISEEIPPEBPSSIEEEVEEVAPPRVP	11197
Qy	114	-----	113
Db	11198	EVIKAPVEAPTVPVKVEAPPAKSKKIPEEKVPVPVQKKEAPPAKVPPEKVPPEKKV	11257
Qy	114	-----	113
Db	11258	LVPKCAVPAPAGRTVLEEKSVAFRQEVVVVKERLELVGEAEVVEEIPBEEFHEVSEYF	11317
Qy	114	-----	113
Db	11318	BEGEPHEVEEFIKLEOHRVBEHRVKVHRVIEVFEABEVEFVKAPKPGPSEKII	11377
Qy	114	-----	113
Db	11378	PPKKPPTKVPRKBEPPAKVPEPKKIIVVEEKVRVPEEPRVPPTKVDLPPKEVVPPEKKV	11437
Qy	114	-----	113
Db	11438	PVPPAKPEAPPKVPPEAPKEVVPPEKKVPVPPPKKPEVPTIKVPEVPKAAVPEKKVPEAI	11497
Qy	114	-----	113
Db	11498	PPKESPPPEVPEAPKEVVPPEKKVPAAPPKKPEVTPVKVPEAPKEVVPEKKVPVPPPKP	11557
Qy	114	-----	113
Db	11558	EVPTKVPDEPKVAPPEKKVPEAITPPKPESPPEPFEEPEBEVALUEBPPEAVEEPEPAAP	11617
Qy	114	-----	113

Db 7238 IRPGNYTITCVGNTPHLRIILKVGKDSGOYTQCATNDVGDKMCQAOLSVKRPFRVKKL 7297  
QY 90 -----LHQDW----- 94  
Db 7298 EASKVAKOGESIQLECKISGPEIKVSWFRNDSELHESWKYNMFSINVSALLTINEASAE 7357  
QY 95 -----||: |----- 94  
Db 7358 DSGDYICEAHNGVGASCSTALTVRKPPVFTQKPSVGALKGSDVILQCEISGTPPFVW 7417  
QY 95 ----- 94  
Db 7418 WVKDRKQVRNKKFKITSKHFDTSUHLNLEASDVGEYHCKATNEVGSDTSCSVKFKEP 7477  
QY 95 -----||: |----- 94  
Db 7478 PRFVKKLSDTSLIGDAVELRAIVEGFQPISVVWLKDRGEVIRESENTRISFIDNIATLQ 7537  
QY 95 ----- 94  
Db 7538 LGSPEASNGKYICQIKNDAGMRECSAVLTVLEPARIIEKPEPMTVTTCGNPFALCVVVG 7597  
QY 95 ----- 94  
Db 7598 TPESAKWFKDGRSADSCKHHITFINKVASLKIPCAEMSKGLYSFEVKNSVGKSNCTV 7657  
QY 95 ----- 94  
Db 7658 SVHVSDRIVPPSFIRKLKDVNALGASVVLGCRVSGSAPISVGWFQDGNIEVSGPKQSS 7717  
QY 95 ----- 94  
Db 7718 FSENVCTLNLSLEPDTGIYTCVAANVAGSDECSAVLTVOEPSPFEQTPDSVEVLPQMS 7777  
QY 95 ----- 94  
Db 7778 LTFTSVIRGTPPFVKVFKGSRVLPGESCNISLEDFVTELEFVQPLESGDYSCLVTN 7837  
QY 95 ----- 94  
Db 7838 DAGSACTHLFVKEPATVFKRLADFSVETGSPVILEATYTGTPISVSWIKDEYILSQS 7897  
QY 95 ----- 94  
Db 7898 ERCSTMTKSTILEILESTIEDYAYQSCLINEAGODICEALVSVLEPPYFIEPLEHVE 7957  
QY 95 ----- 94  
Db 7958 AVIGEPATLQCKDGTPEIRISWYKEHTKLSAPAYKMQFKNNVASLVINKVDHSDVGEY 8017  
QY 95 -----LNKGE----- 99  
Db 8018 SCKADNSVCAVASSAVLVIKERKLPPFFARKLKVHETLGFVPAECPGINSSEPLQVSWY 8077  
QY 100 -----||: |----- 99  
Db 8078 KGVLLKDDANLQTSFVHNATLQILOQDSHIGQYNCSASNPLGTASSAKLILSEHEV 8137  
QY 100 ----- 99  
Db 8138 PPFDDLKPVSDVLALESSTFKCHVTGTAPIKITWAKDNREIRPGNGYKMTLVENTATLT 8197  
QY 100 ----- 99  
Db 8198 VLKVGKGDAQYTCVASNLAGKSCSAHLGVQEPFRFICKLEPSRIVKODEFTRYECKIG 8257  
QY 100 ----- 99  
Db 8258 GSPEIKVLWKDETEIQESSKFRMSFVDSVAVLEMHNLSVEDSGDYTCEAHNAAGSASS 8317  
QY 100 ----- 99  
Db 8318 TSLKVKEPPIFRKKPHPIETLKGADVHLECELGQTPPFPHVSWYKDKRELRSCKYKIMSE 8377

QY 100 ----- 99  
Db 8378 NFLTSIHILNVDAAADIGEYQCKATNDVGDTCVGSIALKAPPRFYKKLSLSDISTVVGKEVQ 8437  
QY 100 ----- 99  
Db 8438 LQTTIEGAEPISVWVFKDKGEIVRESNIIWISYSENIATLQFSRVEPANAGKYTCQIKND 8497  
QY 100 ----- 99  
Db 8498 AGMQECFATLSVLEPATIVERPESIKVTTGDTCTLECTVAGPELSTKWFKDGKELTSDN 8557  
QY 100 ----- 99  
Db 8558 KYKISFFNKVSGKLIINVAPSDSGVYSFEVQNPVGKDSCTASLQVSDRTVPSPFTRKLKE 8617  
QY 100 ----- 99  
Db 8618 TNLGSSVWMECKYVGGPPISSVWFHEGNEISSGRKYQTTLTNTCALTVMNLEESDGG 8677  
QY 100 ----- 99  
Db 8678 DYTCTIATNWAGSDECSAPLTVREPPSFVQKDPDMVDLTGTNTVFTTSIVKGTPPFSVSWFK 8737  
QY 100 ----- 99  
Db 8738 GSSELVPGDRCNVLSLEDSVAELELFDVDTSQSGEYTCIVSNEAGKASCTTHLY IKAPAKF 8797  
QY 100 ----- 99  
Db 8798 VKRLNDYSIEKGKPLILEGTFGTGTPPISVTKWKKNGINVTSPQRNITTTKSAILEIPSS 8857  
QY 100 ----- 99  
Db 8858 TVEDAGQNCYNIENASGKSCSAQIILILEPPYFVKQLEPVPKVSVDGSASLQCLAGTPEI 8917  
QY 100 ----- 99  
Db 8918 GVSWYKGTCLRPRTTYTKMHFRNNVATLVFNQVDINDSGEYICKAENSVEGVSASTFLTIV 8977  
QY 100 ----- 99  
Db 8978 QEQKLPPSFRQLRDVQETVGLPVVFDCAISGSEPISSVSWYKDGKPLKDSNPNVQSFLDN 9037  
QY 100 ----- 99  
Db 9038 TATLNIKTDRSLAGQYSCATNPIGSASSARLILTEGKNPPFFDIRLAPVDVAVGESA 9097  
QY 100 ----- 99  
Db 9098 DFECHVTGTQPIKVSWAKDSREIRSGGKYISYLENSAHLTVLKVDKDGSGQYTCYAVNE 9157  
QY 100 ----- 99  
Db 9158 VKDSCTAQLNIKERLIPPSTFKRLSETVEETEGNSFKLEGRVAGSQPITVAVYKNNIEI 9217  
QY 100 -----YCKVSN----- 106  
Db 9218 QPSTNCEITFRNNTLVLOVRKAGMNDAGLYTCKVSNDSAGSALCTSSIVIKEPKKPPVFDQ 9277  
QY 107 -----||: |----- 106  
Db 9278 HLTPTVSEGEYVOLSCHVQSGSEPIRIQWLKAGREIKPSDRCSFSFASCTAVLELRDVAK 9337  
QY 107 ----- 106  
Db 9338 ADSGDYVCKASNVAGSDTTKSKVTIKDKPAVAPATKAAVDGRLFFVSEPOIRVVEKTT 9397  
QY 107 ----- 106  
Db 9398 ATFTIAKVGDDPIPNNVWTKGKWRQLNQGRVFIHOKGDEAKLEIRDTTKTDSGLRYCVAF 9457

QY	81	-----	80
Db	5078	KEIADRYRIAFVEGTASLEIIRVDMNDAGNFTCRATNSVSGKSSGALLIYQPPSFVT	5137
QY	81	-----	80
Db	5138	KPGSKDVLPGSAVCLAKSTFOGSTPLTIRWPKGNKELVSGGSCYITKEALESLELYIVKT	5197
QY	81	-----	80
Db	5198	SDSGTYTCKVSNVAGGVECSANLFEVKEPATFVEKLEPSQLLKKGDATQLACKVTGTPPIK	5257
QY	81	-----	80
Db	5258	ITWFANDREIKESKHRMSFVESTAVLRLTDVGIEDSGEYMCEAONEAGSDHCSSIVIVK	5317
QY	81	-----	80
Db	5318	ESPYFTKEFKPIEVLKEYDYMLLAEVAGTTPPEITWFKONTILRSGRKYKFTIQDHLVSL	5377
QY	81	-----	80
Db	5378	QILKFAADAGEYQCRVTNEVGSSICSARVTLREPPSFYKKBESTSSLRGGTAFOATLK	5437
QY	81	-----	80
Db	5438	GSLPITVTWLKDSDEITEDDNIRMTFENNVAASYLSGIEVKHDKYVCOAKNDAGTORCS	5497
QY	81	-----	80
Db	5498	ALLSVKREPATITEAVSIDVTQGDPATLOVKFSGTKEITAKWFKDQELTLGSKYKISVT	5557
QY	81	-----	80
Db	5558	DTVSIILKIISTEKDSGEYTFEVQNDVGRSSCKARINVLDLIIPSTFKLKMDSIKGS	5617
QY	81	-----	80
Db	5618	FIDLECIVAGSHPISIOWFKDQEIISASEKYKFSFHDNTAFLEISQLEGTDSTYTCSAT	5677
QY	81	-----	80
Db	5678	NKAGHNQCSGHLRVKEPPYFVEKXPQODVNPTRVOLKALVGTAPTIKWFKDNKELHS	5737
QY	81	-----	80
Db	5738	GAARVWKDDTSTSLFAAKATDSGTYYICQLSNDVGTATSKATLFVKEPPQFIKKPSV	5797
QY	81	-----	80
Db	5798	LVLNQGOSTTFECQITGTPKIRVSWYLDGNEITAIQKHGISFIDGLATFQISGARVENS	5857
QY	81	-----	80
Db	5858	TYVCEARNDAGTASCIELVKKEPPTFIRELKPVEVVKYSDVECEVTGTPPEVTWLK	5917
QY	81	-----	80
Db	5918	NNREIRSSKYYLTDRVSVFNHLITKCDPSDTGEYQCIIVSNEGGSCSTRVALKEPPSF	5977
QY	81	-----	80
Db	5978	IKKIENITVLKSSATFQSTVAGSPISITWLKDDQILDEDDNNVYISFVDSVATLIQIRSV	6037
QY	81	-----	80
Db	6038	DNGHSGRYTQAKNESGVERCYAFLLVQBPQAIWEKAKSVDVTEKDPMTLECVVAGTPEL	6097
QY	81	-----	80
Db	6098	KVKWLKDGKQIVPSRYFSMFENNVAIFRIQSVMKODSQYTFKVENDFGSSCDAYLRV	6157
QY	81	-----	84

Db	6158	LDQNIIPPFTKLTLMKMDKVLGSSIIHMECKVSGSLPISAOQWFKDCKEISTSAKYRLVCHER	6217
QY	85	-----	84
Db	6218	SVSLEYNNLEEDTANYTCKVSNVAGDDACSGILTVPKPSFLVKPQROQAIPDSTVEFK	6277
QY	85	-----	84
Db	6278	AILKGTTPPKIKWFKDDVELVSGPKCFIGLEGSTSFNLVSYVDASKTGQYTCHTVNDVGS	6337
QY	85	-----	84
Db	6338	DSCTTMLLVTEPPKFKVVKLEASKIVKAGDSSRLCKIAGSPEIRVVWFRNEHELPSADKY	6397
QY	85	-----	84
Db	6398	RMTFIDSVAVIOMNNLSTEDSGDFICEAQNPAGSTSCSTKVIVKEPPVFSSFPPIVETLK	6457
QY	85	-----	84
Db	6458	NAEVSLECELSGTPPEFVWYKDKROLRSKKYKIAKSNFHTSIHILNVOTSDIGEYHCK	6517
QY	85	-----	84
Db	6518	AQNEVGSDTCVCTVKLKEPPRFVSKLNSLTVVAGEPAELQASIEGAQPIFVQWLKEKEEV	6577
QY	85	-----	84
Db	6578	IRESENIRITFVENVATLOFAKAEPANAGKYICQIKNDGMEENMATLMVLEPAVIVEKA	6637
QY	85	-----	84
Db	6638	GPMTVTVGETCTLECKVAGTPELSVENYKDGKLLTSQXHKHFSFYKNKISSRLILSVRQD	6697
QY	85	-----	84
Db	6698	AGTYTFQVQNNVCKSSCTAVVDVSDRAVPPSTFRLKNTGGVLGASCILECKVAGSSPIS	6757
QY	85	-----	SVLTV- 89 : : : : : : : : : :
Db	6758	VAMFHEKTIKIVGAKYQTTTFSDNVCTQLNLSLSDSDMGNYTCVAANVAGSDECAVLTVQ	6817
QY	90	-----	89
Db	6818	EPPSFVKEPEPLEVLPKKNVTFTSVIRGTPPKVNNFRGARELVKGDRONIYPEDTVAEL	6877
QY	90	-----	89
Db	6878	ELFNIDISQSGEYTCVVSNAGOASCTTRLFVKPEAAFLKRLSDHSVEPKSILLESTYT	6937
QY	90	-----	89
Db	6938	GTLPISTVWKDKGNITTEKCNIVTTEKTCILEILNSTKRDAGQYSCIEENAGRDVCG	6997
QY	90	-----	89
Db	6998	ALVSTLEPPYFVTELEPLEAAVGDVSLQCVAGTPEITVSWYKGTCLKLRPTPEYRTYFT	7057
QY	90	-----	89
Db	7058	NNVATLVFNKVNINDSGEYTCKAENSIGTASSKTVPRIQERQLPPSFARQLKDIEQTVGL	7117
QY	90	-----	89
Db	7118	PVTLTCLRLNGSAPIQVCWYRDRGVLRRDDENLOTSFVDNVATILKILQTDLSHSGQYSCSAS	7177
QY	90	-----	89
Db	7178	NPLGTASSARLTAREPKKSPFFDIKPSIDVIAGESADFECHVTGAQPMRITWSKDNKE	7237
QY	90	-----	89

Db 2858 PSDAGEYTAVGQLECKAKLFVETLHITKTMKNIEVPETKTASFCEVSHFNVPMSWLKN 2917  
QY 74 ----- 73  
Db 2918 GVEIEMSEKFIKVOGKLHQLIIMNTSTEDSAEYTFVCGNDQVSATLVTPTIMTSMKLD 2977  
QY 74 ----- 73  
Db 2978 INAEKDTITFEVTVNYEGISYKWLKNGVEIKSTDKCQMRKTLHSLNIRNVHFGDAAD 3037  
QY 74 ----- 73  
Db 3038 YTFVAGKATSTATLYVEARHIEFRKHIDIKVLEKKRAMFECEVSEPDITVOMMKDDQEL 3097  
QY 74 ----- 73  
Db 3098 QITDRIKIQEKYVHRLIIPSTRMSDAGKYTVVAGNVSTAKLFVEGRDVRIRSIKKEVQ 3157  
QY 74 ----- 73  
Db 3158 VIEKQAVVEFVNEDVDHAWYKDGIEINFQVQERHKYVVERRIHRMFISETROSADAGE 3217  
QY 74 -----BEQ 76  
Db 3218 YTFVAGNRSSVTLVNAPEPPQVLQELQPVTVQSGKPARFCAVTSGRPPKISWYKBEQ 3277  
QY 77 YNST----- 80  
Db 3278 LLSTGFKCKFLHDGQEYTLILLIEAFPDEAAVYTCENKDYGVATTSASLSVEVPEVSPD 3337  
QY 81 ----- 80  
Db 3338 QEMPVYPAIITPLOTVTVTSEGQPARFCRVSGTDLKVSWSKDKKIKRPSFRMTQPED 3397  
QY 81 ----- 80  
Db 3398 TYOLEIAEAYPEDEGTYTFVNASNAVQVSSANLSLEAPESILHERIEQEIEMEMKEFPSS 3457  
QY 81 ----- 80  
Db 3458 SFLSAEEGLHSAELQSKINETLELLESPPVYSTKFDSEKGTGPIFIKEVSNADISMG 3517  
QY 81 ----- 80  
Db 3518 DVATLSVTVIGIPKPIQWFFNGVLLTPSADYKFVFDGDHSLIILFTKLEDEGEYTCMA 3577  
QY 81 ----- 80  
Db 3578 SNDYKGTICSAYLKINSKGEHGDTESETSAVAKSLEKLGPCPPHFLKELKPIRCAQGLP 3637  
QY 81 ----- 80  
Db 3638 AIFETVVGEPAPTVTWEKENQLCTSVYTIHNPNGSGTFIVNDQREDSGLYICKAE 3697  
QY 81 ----- 80  
Db 3698 NMLGESTCAEALLVLEDDMTDTPCKAKSTPEAPDFPQTLKGPVAVALDSEQEIATF 3757  
QY 81 ----- 80  
Db 3758 VKDTILKAALITEENQQLSYEHIAKANELSSQLPLGAQELQSLQEDKLTPESTREFLCI 3817  
QY 81 ----- 80  
Db 3818 NGSIHFQPLKPSNQLQIVOSOKTFSKEGILMPEEPETOAVLSDTEKIFPSAMSIEQI 3877  
QY 81 ----- 80  
Db 3878 NSLTVEPLKTLAEPEGNYPOSSIEPPMHSYLTSVAEVLSPKKTVDNREQRVTLOK 3937  
QY 81 ----- 80  
Db 3938 QEAQSALILSQSLAECHVESLQSPDVMISQVNYEPLVPSEHSECTEGGKILIESANPLENA 3997

QY 81 ----- 80  
Db 3998 QODSAVRIEEGKSLRFPPLALEEKQVLLKEHSDNVVMPDQILLESKREPVAIKKVQEVQ 4057  
QY 81 ----- 80  
Db 4058 RDLLESKILLSGIPPEQRNLKIQICRALQAAVASEQPGFSEWLRNRIEKVEAVNITQ 4117  
QY 81 ----- 80  
Db 4118 EPRHMCWYLVTSAKSVTEEVTTIIIEDVDPQMANLKMELRDALCAIYIEEIDILTAEGPR 4177  
QY 81 ----- 80  
Db 4178 IQOGAKTSLQEEWDSFSGSQKVEPITEPEVESKYLISPEEVSFVNVQSRVKYLDATPVTK 4237  
QY 81 ----- 80  
Db 4238 GVASAVSDEKODESLKPFSEKESSESSEGTVEEVATVKIQEAGGFIKEDGPMIHTPLVD 4297  
QY 81 ----- 80  
Db 4298 TVSEEGDIVHLTTSITNAKEVNWYFENKLVPSDEKFKCLQDQNTYTLVIDKVNTEHDQGE 4357  
QY 81 ----- 80  
Db 4358 YVCEALNDSGTKTSAKLTVVKRAAPVIKRIEPLLEVALGHLAKFTCEIQSAPNVRFQWF 4417  
QY 81 ----- 80  
Db 4418 KAGREIYESDKSIRSSKYISSLEILRTQVDCGEYTCASNEYGSVSTATLTVTEAYP 4477  
QY 81 ----- 80  
Db 4478 PTFLSRPKSLTTFVCKAAKFICTVTGTPVIETIWKDGAALSPSPNWKISDAENKHILEL 4537  
QY 81 ----- 80  
Db 4538 SNLTIDRGVYCKASKNFGADICQAEI I IDKPHFIKELEPVQSAINKKHLEQOVDED 4597  
QY 81 ----- 80  
Db 4598 RKTVTWSKDGOKLPPGDKYKICFEDKIATLEIPLAKLSDGTYVCTASNEAGSSCSAT 4657  
QY 81 ----- 80  
Db 4658 YTVREPPSFVKVDPYSYLMPLGESARLHCKLKGSPVIQVTFKNNKELSESNTVRMYFVN 4717  
QY 81 ----- 80  
Db 4718 SEAILDITDKVEDSGSYSCAVNDVGDSCSTEIVIKPEPPFIKTLEPADIVRGTNALL 4777  
QY 81 ----- 80  
Db 4778 QCEVSGTGPFEISWFKDKKQIRSKKYRFLFSOKLSVCLCIFSNADGVEYECVVANEVG 4837  
QY 81 ----- 80  
Db 4838 KGCMAATHLLKEPPTFVKVDDLIALGGOTVTLQAARGSEPISTVWMMKGQEVIREDKGI 4897  
QY 81 ----- 80  
Db 4898 KMSFNGVAVLIIPDVQISFGGKYTCLAENAGSOTSVGELIVKEPAKIERAELIQVTA 4957  
QY 81 ----- 80  
Db 4958 GDPATLEVTVAGTPELKPWKYKGRPLVASKKYRISFKNNVAQLKFYSAEHLDSGOYTE 5017  
QY 81 ----- 80  
Db 5018 ISNEVGSSCETTFTVLDRDIAFFTTPLRNVDSVVVNGTCRLDCKIAGSLPMRVSFKDQ 5077

QY 10 ----- 9  
Db 698 KAEAVATVAADQARVREPREGHLEESYAQQTILEYGYKERISAAKVAEPQRPASEP 757  
QY 10 ----- 9  
Db 758 HVPKAVPRVIOAPSETHIKTTDQGMHISQIKKTTDLTTERLVHVKRRPTASPHET 817  
QY 10 ----- 9  
Db 818 VSKISVPKTEHGEYASIAISATLOKELSATSSAOKITKSVKAPTVPKSETRVRAEPTP 877  
QY 10 -----CPAP-- 13  
Db 878 LPQFPADPTDYKSEAGVEVKVCVSGITGVREERFEVLHGREAKVTETARVPAPVE 937  
QY 14 -----ELGQPSVFL-----FP----- 25  
Db 938 IPVTPTLVSLGNVTVIEGESVTLCHISGVPSTVTWYREDYQIESSIDFOITFQSGI 997  
QY 26 ----- 25  
Db 998 ARLMIREAFADSGRTCSAVNEAGTVSTCYLAVOVSEEFKETTAVTEKFTTEKRFV 1057  
QY 26 -----PKP----- 28  
Db 1058 ESRDVVMTDTSLTEQAGPEPAAPYFITKPVVQKLVEGVSVFGCVGNPKPHYWK 1117  
QY 29 ----- 28  
Db 1118 SGVPLTGYRYKVSYNKQGEKLVISMTFADDAGEYTIIVRNKHGETSASASLLEADY 1177  
QY 29 -----KDTLM-- 33  
Db 1178 ELLMKSQEMLYQTVTAFQEPKVGCTAPGFVYSEYEKEQEALIRKKMAKDTVVVR 1237  
QY 34 ----- 33  
Db 1238 TVVEDQEFHISFEERLIKEIYRIKTTLEELLEDEGEKMAVDISESEAVESGFDLRI 1297  
QY 34 ----- 33  
Db 1298 KNYRILEGMGVTHCKMSGYPLPKIAYKDGRIKHGERYQMDFLQDGRASLRIPVYLPE 1357  
QY 34 -----ISRTP-- 38  
Db 1358 DEGIYTFASNKGNACSGKLYVEPAAPLGAPTYIPTLEPVSIRSLSPRSVSRPIRM 1417  
QY 39 ----- 38  
Db 1418 SPARMSPARMSPARMSPGRRLEETDESQERLYKPVFVKPVSPKCLEGGQTARFD 1477  
QY 39 ----- 38  
Db 1478 LKVVGRMPETFWHDGOQIVNDYTHKVVIKEDGTOSLIIVPATPSDGEWTVVAONRAG 1537  
QY 39 ----- 38  
Db 1538 RSSISVILTVEAVEHQVKPMFVEKLNKVNKEGSQLEMKVRATGNPNPDIVLKNSDIIV 1597  
QY 39 ----- 38  
Db 1598 PHYPKIRIEGTGEAALKIDSTVSQDSAWYTATAINKAGRODTRCKVNVEFEFAPEPE 1657  
QY 39 ----- 38  
Db 1658 RKLIIIPRGTYRAETAPEPLHLRYGOEWEEGLDYDKEKQKQPFKKKLTSLRKR 1717  
QY 39 -----EUTC 42  
Db 1718 GPAHFECRLTPIGDTMVVWVWHLHGKPLEAANRLRMINEFGYCSLDYGVAYSRDSGIITC 1777  
QY 43 ----- 42

Db 1778 RATNKYGTDHTSATLIVKDEKSLVESQLPEGRKGLQRIEELERMAHEGALTGVTTDQKE 1837  
QY 43 ----- 42  
Db 1838 KOKPDIIVLYPEPVRVLEGETARFCRVRTGYPOPKNWYLMGQLIRKSRFRVYDGIHYL 1897  
QY 43 ----- 42  
Db 1898 DIVDCKSYDTGEVVKVTAENPEGVIEHKVLEIQQOEDFRSLVRRAPRPFHFVHEPGKL 1957  
QY 43 ----- 42  
Db 1958 QPEVQKVRPDVTETKEVVKLKRABRITHKEKVPESSEELRSFKRRTBEGYVEAITAVE 2017  
QY 43 ----- 42  
Db 2018 LKSRKKDESYEELLRRKTDELHMTKELTEEEKKALAEEGKITPTFKPKIELSPSMEA 2077  
QY 43 -----V 43  
Db 2078 PKIFERIQSQTVGQGSDAHFRVRVCKPDPECEWYKNGVKIERSDRIYWNPEDNVCELY 2137  
QY 44 VVDVSHED----- 51  
Db 2138 IRDVTAEADSASIMVKAINIAGETSSHAFLLVQAKQLITFTQELQDVVAKEDTMTATPECE 2197  
QY 52 ---PEVKFNWYDGVGVH-----NAKT- 70  
Db 2198 TSEPFVKVWKYKDGMEVHEGDKYRMHSRDKRVHFLSILTIDTSDAEDYSCVLVEDENVKTT 2257  
QY 71 ----- 70  
Db 2258 AKLIVEGAVVEFVKELQDIEVPESYSGELECIYSPENIEGKWYHNDVELKSNKYTITSR 2317  
QY 71 -----KPR----- 73  
Db 2318 RGRQNLTVKDVTKEDQGEYSFVIDGKKTCKLKMKPRPAILQGLSDQKVCEGDIVQLEV 2377  
QY 74 ----- 73  
Db 2378 KVSLESVEGVMKDGQEVQPSDRHVIVIDKQSHMLIEDMTKEDAGNYSFTIPALGLSTS 2437  
QY 74 ----- 73  
Db 2438 GRVSVYSVDVITPLKDVNVIEGTKAVLECKVSVPDVTSVKWYLNDEQIKPDDRVQAIKVG 2497  
QY 74 ----- 73  
Db 2498 TKQRLVNRTHASDEGPKLIVGRVETNCNLSVEKIKIIRGLDLCTCTETQNVVFEVELS 2557  
QY 74 ----- 73  
Db 2558 HSGIDVLWNFKDKEIKPSSKYKIEAHGKIYKLTVLNMKDDGKYTYFAGENITSGKLTV 2617  
QY 74 ----- 73  
Db 2618 AGGAIKSKPLTDQTVAESQEAVFCEVANPDSKGEWLBDGKHLPLTNNIRSESDGHRRLI 2677  
QY 74 ----- 73  
Db 2678 IAATKDDIGEYTKVATSKTSAKLKVEAVKIKTKLNLTVTETQDAVFTVELTHPNVKG 2737  
QY 74 ----- 73  
Db 2738 VQMIKNGVVLESNEYAISVKGTIYSLRIKNCIAIVDESIVGFRLGASARLHVETVKI 2797  
QY 74 ----- 73  
Db 2798 IKKPKDVTALENATVAFEVSVSHDTPVPKWFHKNVEIKPSDKHRLVSRKRVHKLMLQNIS 2857  
QY 74 ----- 73

RA Strausberg R.;  
RL Submitted (JUL-2001) to the EMBL/GenBank/DBJ databases.  
DR EMBL; BC010327; AAH10327.1; -  
DR MGI; MGI:2144967; AU044919.  
DR InterPro; IPR000345; CytC\_heme\_bind.  
DR InterPro; IPR003006; I9\_MHC.  
DR Pfam; PF00047; I9; 3.  
DR PROSITE; PS00190; CYTOCHROME\_C; UNKNOWN\_1.  
DR PROSITE; PS00290; I9\_MHC; UNKNOWN\_1.  
KW Hypothetical protein.  
SQ SEQUENCE 473 AA; 51946 MW; CF625F008932AF12 CRC64;

Query Match 59.0%; Score 787; DB 11; Length 473;  
Best Local Similarity 50.0%; Pred. No. 3.3e-23;  
Matches 143; Conservative 34; Mismatches 47; Indels 62; Gaps 2;

QY 5 HT-----CP 8  
DB 188 HTPFALLQSLGYTMSSVTPSPSTVTCVAHPASSTTVDDKKLEPSGPISTINPCP 247  
QY 9 P-----CPAPELLGGPSVFLPDKPKDLMISRTPEVTCVVVDVSHEDPEVKFNVDG 62  
DB 248 PCKECHCKAPNLEGSPSFIFFPNKIDVLMISLTPKVCVVVDVSEDDPDVQISFEVNN 307  
QY 63 VEVNNAKTPREQYNSTYRVSVLVTLVHODWLNKKEYCKVSKNALKAPAPIERTISKAG 122  
DB 308 VEVHTAQTOHREDYNSTIRVSALPIQHDWMSGKEFKCKVKNKDLPSPIERTISKIKG 367  
QY 123 QPREPOVYTLPPSRDELTKNOVSLTCLVKGFPSPDIAVEWESNGQPENNYKTPPPVLDS 182  
DB 368 LVRAPOVYILPPPAEQLSRKDVSCLTVGVNFGDISVETWSNGHTEENYKDPAPVLDS 427  
QY 183 GSFFLYSKLTVDKSRVQOGNVFSCSVYMHKALHNHYTKLSLSPGK 228  
DB 428 GSYFYSKLDIKTSKEKTDSCVNRHGLKNYILKXISRSPGK 473

RESULT 13  
Q8WZ42  
ID Q8WZ42 PRELIMINARY; PRT; 34350 AA.  
AC Q8WZ42;  
DT 01-MAR-2002 (TrEMBLrel. 20, Created)  
DT 01-MAR-2002 (TrEMBLrel. 20, Last sequence update)  
DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)  
DE titin.  
GN TTN.  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
OX NCBI\_TaxID=9606;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=20309627; PubMed=10850961;  
RA Freiburg A., Trombitas K., Hell W., Cazorla O., Fougereousse F.,  
RA Centner H., Kolmerer B., Witt C., Beckmann J.S., Gregorio C.C.,  
RA Granzier H., Labeit S.;  
RT "Series of exon-skipping events in the elastic spring region of titin  
RT as the structural basis for myofibrillar elastic diversity.";  
RL Circ. Res. 86:1114-1121(2000).  
RN [2]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=21573839; PubMed=11717165;  
RA Bang M.L., Centner T., Fornoff F., Geach A.J., Gotthardt M.,  
RA McNabb M., Witt C.C., Labeit D., Gregorio C.C., Granzier H.,  
RA Labeit S.;  
RT "The complete gene sequence of titin, expression of an unusual -700  
RT 'Ada titin isoform and its interaction with obscurin identify a novel  
RT 2-line to i-band linking system.";  
RL Circ. Res. 89:1065-1072(2001).  
DR EMBL; AJ277892; CAD12456.1; -  
DR InterPro; IPR000282; Cytok\_receptor\_2.  
DR InterPro; IPR000719; Euk\_pkinase.  
DR InterPro; IPR000577; FGKY\_kin.

DR InterPro; IPR003961; FN\_III.  
DR InterPro; IPR001092; HLH\_basic.  
DR InterPro; IPR003599; I9.  
DR InterPro; IPR003598; I9\_c2.  
DR InterPro; IPR003006; I9\_MHC.  
DR InterPro; IPR003596; I9\_v.  
DR InterPro; IPR002016; Peroxidase.  
DR InterPro; IPR004168; PPAK\_motif.  
DR InterPro; IPR002290; Ser\_thr\_pkinase.  
DR InterPro; IPR001245; Tyr\_pkinase.  
DR Pfam; PF00041; fn3; 132.  
DR Pfam; PF00047; I9; 146.  
DR Pfam; PF00069; pkinase; 1.  
DR Pfam; PF02818; PPAK; 53.  
DR ProDom; PD000001; Euk\_pkinase; 1.  
DR SMART; SM00060; FN3; 133.  
DR SMART; SM00409; I9; 167.  
DR SMART; SM00408; IGC2; 148.  
DR SMART; SM00406; IGV; 23.  
DR SMART; SM00220; S\_TKC; 1.  
DR SMART; SM00219; TYKc; 1.  
DR PROSITE; PS00933; FGKY\_KINASES\_1; UNKNOWN\_1.  
DR PROSITE; PS00038; HELIX\_LOOP\_HELIX; UNKNOWN\_1.  
DR PROSITE; PS00290; I9\_MHC; UNKNOWN\_1.  
DR PROSITE; PS00435; PEROXIDASE\_1; UNKNOWN\_1.  
DR PROSITE; PS00011; PROTEIN\_KINASE\_DOM; 1.  
DR PROSITE; PS00109; PROTEIN\_KINASE\_TYR; UNKNOWN\_1.  
SQ SEQUENCE 34350 AA; 3816262 MW; 5B1120058A7CE58A CRC64;

Query Match 36.4%; Score 486; DB 4; Length 34350;  
Best Local Similarity 0.5%; Pred. No. 9.7e-05;  
Matches 160; Conservative 43; Mismatches 40; Indels 33574; Gaps 45;

QY 3 KTHTCPP-----9  
DB 98 KAETAPPNFVRLQSMVTGROGQVRLQVRVTGPTPVVKFYRDGAETQSSLDFOISEGD 157  
QY 10 -----9  
DB 158 LYSLLIAEAYPEDSGIYSVNNATNSVGRATSTAELLVQGEVEVPKTKTIVSTAQISESR 217  
QY 10 -----9  
DB 218 QTRIEKIEAHFDARSIAIVEMIDGAAGQOLPHKTPHRIIPKPKSRPTPPSTAARQAOL 277  
QY 10 -----9  
DB 278 AROQSPSPINHSPSPVHRVRAPTPPSVRSVSPAARISTSPRSPLLMRKQASTVAT 337  
QY 10 -----9  
DB 338 GPEVPPPKQEGYVASSSEAEARETTTLTTSTQIRTEERWEGYGVQVQVTSIGAAGAAAS 397  
QY 10 -----9  
DB 398 VSASAYAAEAVATGAKVQKQADKSAAVATVVAADMARVREPVISAVEQTAQRTTTA 457  
QY 10 -----9  
DB 458 VHIQPAQEQVRKEAEKTAIVTKVVVAADKAKEQELKSRKEVITTKQEQMHVTHEQIRKET 517  
QY 10 -----9  
DB 518 EKTVPKVVISAAKAEQETRISEITKKQKVQTEAIRQETETITAASWVVVATAKSTKL 577  
QY 10 -----9  
DB 578 ETVPGAQETTTQDDQMHLSYKIMKTRKTVVVKVIVATPKVQDLVSRGREGITTKR 637  
QY 10 -----9  
DB 638 EQVQITQEKMRKEAEKTAISTIAVATAKAEQETILTRETMTATRQEQIQVTHGKVDVGK 697

```
DR SMART: SM00406; IGV; 1.
DR SMART: SM00410; IG-like; 1.
DR PROSITE: PS00290; IG_MHC; UNKNOWN_1.
SQ SEQUENCE 473 AA; 52449 MW; BE9889B7986DA155 CRC64;

Query Match
Best Local Similarity 60.8%; Score 811; DB 11; Length 473;
Matches 148; Conservative 30; Mismatches 46; Indels 56; Gaps 2;

QY 5 HT-----CPP- 9
DQ 194 HTFAVLQSDLYTLSSVTVTSSWPSQITCNVAHPASSTKVDKKEIPRGFTIKPCPPC 253
QY 10 -CPAPPELLGGPSVFLFPPKPKDTLMISRTPEVTCVVVDVSHEDPEVKFNWYVDGVEVHNA 68
DQ 254 KCPAPNLLGGPSVFLFPPKPKDTLMISRTPEVTCVVVDVSHEDPEVKFNWYVDGVEVHNA 313
QY 69 KTKPREQYNSTYRVVSVLTVLHQDWLNGKEYKCKVSNKALPAPIEKTIISKAKGPREPQ 128
DQ 314 QTOFHREDYNSTLRVVSALPQHQQDMWKGKFKCKVNNKALPAPIERTISKPKGSVRAPQ 373
QY 129 VYTLPPSRDELTKNQVSLTCLVKGFYPSDIAVEWESNGQPENNYKTTPPVLDSDGSFPLY 188
DQ 374 VYVLPPEEEMTKKQVLTCTMVTDFMPEDIVYEWNTNGKTELNYKNTPEVLDSDGSFPLY 433
QY 189 SKLTVDKSRWQGVNFCSSVMHEALHNHYTOKLSLSPGK 228
DQ 434 SKLRVERKKNWERNYSYCSVVHGLNHHHTKFSFRTPGK 473

RESULT 10
Q99L31 ID Q99L31 PRELIMINARY; PRT; 468 AA.
AC Q99L31;
DT 01-JUN-2001 (Tremblrel. 17, Created)
DR EMBL; BC003878; AA03878.1;
DR HSP; P01842; 7FAB.
DR InterPro; IPR003599; IG.
DR InterPro; IPR003597; IG-cl.
DR InterPro; IPR003600; IG-like.
DR InterPro; IPR003006; IG_MHC.
DR InterPro; IPR003596; IG_v.
DR Pfam; PF00047; ig; 3.
DR SMART; SM00409; IG; 2.
DR SMART; SM00407; IGcl; 3.
DR SMART; SM00406; IGV; 1.
DR SMART; SM00410; IG-like; 1.
DR PROSITE; PS00290; IG_MHC; UNKNOWN_1.
SQ SEQUENCE 468 AA; 51661 MW; 96352328B3332ADB CRC64;

Query Match
Best Local Similarity 52.9%; Score 810; DB 11; Length 468;
Matches 148; Conservative 30; Mismatches 46; Indels 56; Gaps 2;

QY 5 HT-----CPP- 9
DQ 189 HTFAVLQSDLYTLSSVTVTSSWPSQITCNVAHPASSTKVDKKEIPRGFTIKPCPPC 248
QY 10 -CPAPPELLGGPSVFLFPPKPKDTLMISRTPEVTCVVVDVSHEDPEVKFNWYVDGVEVHNA 68
DQ 249 KCPAPNLLGGPSVFLFPPKPKDTLMISRTPEVTCVVVDVSHEDPEVKFNWYVDGVEVHNA 308
```

```
QY 69 KTKPREQYNSTYRVVSVLTVLHQDWLNGKEYKCKVSNKALPAPIEKTIISKAKGPREPQ 128
DQ 309 QTOFHREDYNSTLRVVSALPQHQQDMWKGKFKCKVNNKALPAPIERTISKPKGSVRAPQ 368
QY 129 VYTLPPSRDELTKNQVSLTCLVKGFYPSDIAVEWESNGQPENNYKTTPPVLDSDGSFPLY 188
DQ 369 VYVLPPEEEMTKKQVLTCTMVTDFMPEDIVYEWNTNGKTELNYKNTPEVLDSDGSFPLY 428
QY 189 SKLTVDKSRWQGVNFCSSVMHEALHNHYTOKLSLSPGK 228
DQ 429 SKLRVERKKNWERNYSYCSVVHGLNHHHTKFSFRTPGK 468

RESULT 11
Q8R3H6 ID Q8R3H6 PRELIMINARY; PRT; 474 AA.
AC Q8R3H6;
DT 01-JUN-2002 (Tremblrel. 21, Created)
DT 01-JUN-2002 (Tremblrel. 21, Last sequence update)
DT 01-JUN-2002 (Tremblrel. 21, Last annotation update)
DE Hypothetical 51.7 kDa protein.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RA Strausberg R.;
RL Submitted (MAR-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC025447; AA025447.1;
KW Hypothetical protein.
SQ SEQUENCE 474 AA; 51748 MW; 8608B57C6CD2874A CRC64;

Query Match
Best Local Similarity 59.1%; Score 788; DB 11; Length 474;
Matches 143; Conservative 34; Mismatches 47; Indels 62; Gaps 2;

QY 5 HT-----CP 8
DQ 189 HTFPALQSLGYTMSSSVTVPSSTWPSQVTVCSVAHPASSTVVDKLEPSGPISTINPCP 248
QY 9 P-----CPAPPELLGGPSVFLFPPKPKDTLMISRTPEVTCVVVDVSHEDPEVKFNWYVDG 62
DQ 249 PKCECHCKCAPNLEGPSVFIFPPNRIKDLVLMISLTQVTKVTVVVDVSEDDPDVQISWFEVNN 308
QY 63 VEHNATKPREQYNSTYRVVSVLTVLHQDWLNGKEYKCKVSNKALPAPIEKTIISKAKG 122
DQ 309 VEHTAQTQTHREDYNSTIRVVSALPQHQQDMWKGKFKCKVNNKALPAPIERTISKIKG 368
QY 123 QPREPOVYTLPPSRDELTKNQVSLTCLVKGFYPSDIAVEWESNGQPENNYKTTPPVLDSD 182
DQ 369 LVRAPOVYILPPPAEQLSRKDVSLTCLVGVNPGDISVEWTSNGTEENYKDTAPVLDSD 428
QY 183 GSFFLYSKLTVDKSRWQGVNFCSSVMHEALHNHYTOKLSLSPGK 228
DQ 429 GSVFIYKSLDKITKWEKTDSEFCNVRHEGLKNTLTKTISRSPGK 474

RESULT 12
Q91Z05 ID Q91Z05 PRELIMINARY; PRT; 473 AA.
AC Q91Z05;
DT 01-DEC-2001 (Tremblrel. 19, Created)
DT 01-DEC-2001 (Tremblrel. 19, Last sequence update)
DT 01-JUN-2002 (Tremblrel. 21, Last annotation update)
DE Hypothetical 51.9 kDa protein.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
```



```

RT antibody (Mab 7, its light and heavy chains) and construction of a
RT single chain antibody (scFv).";
RL Submitted (MAY-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL: AF152372; AAD40243.1; -.
DR HSP; P01842; 7FAB.
DR MGD; MGI:96446; Igh-4.
DR InterPro; IPR003600; Ig_Like.
DR InterPro; IPR003006; Ig_MHC.
DR InterPro; IPR003596; Ig_V.
DR Pfam; PF00047; Ig; 4.
DR SMART; SM00406; Igv; 1.
DR SMART; SM00410; IG_Like; 2.
DR PROSITE; PS00290; IG_MHC; UNKNOWN_1.
FT NON_TER 1
FT 437
SQ SEQUENCE 437 AA; 48142 MW; 5C3A7BB3EE7D697C CRC64;

Query Match 62.3%; Score 831; DB 11; Length 437;
Best Local Similarity 51.6%; Pred. No. 4e-25;
Matches 143; Conservative 43; Mismatches 35; Indels 56; Gaps 3;
QY 5 HT-----CpP--Cp 11
Db 164 HTFPAVLQSLYLTSSVTVPSTWPSQITCNVAHPASSTKVKDKIYPRDCGCRPCICT 223
QY 12 APELLGGPSVFLPPKPKDITLMISRTPEVTCVVVDVSHEDPEVKFNWYGVGVEVHNATK 71
Db 224 VPEV---SSVFPPKPKDVLITLPKVTCTVVVDISKDDPEVQFSWFVDDVEVHTAQIQ 280
QY 72 PREQYNSTYRVSVLTVLHQLWLGKEYCKVSNKALPAPIETKISKAKGPREPQVYT 131
Db 281 PREQYNSTYRVSVLTVLHQLWLGKEYCKVSNKALPAPIETKISKAKGPREPQVYT 340
QY 132 LPSPRDELTKNOVSLTCLVKGYFSDIAVFNESGNQPNYKTTTPVLDSGCSFFLYSKL 191
Db 341 IPPPEQMAKDKVSLTCTMTDFPEDIYVQWQNGQPAENYKNTQPIMDTDSGFYYSKL 400
QY 192 TVDKSRWQGNVFCSCVWHEALHNHYTQKSLSLSPGK 228
Db 401 NVQKSNWEAGNTFTCSVLHLEGLHNHHTEKLSHSPGK 437

RESULT 8
Q9D8L4
ID Q9D8L4 PRELIMINARY; PRT; 473 AA.
AC Q9D8L4;
DT 01-JUN-2001 (TrEMBLrel. 17, Created)
DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE 1810060009RIK protein.
GN IGH-1 OR 1810060009RIK.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=PANCREAS;
RX MEDLINE=21085660; PubMed=11217851;
RA Kawai J., Shinagawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,
RA Arakawa T., Hara A., Fukunishi Y., Konno H., Adachi J., Fukuda S.,
RA Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamanaka I.,
RA Saito T., Okazaki Y., Gojobori T., Bono H., Kasukawa T., Saito R.,
RA Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,
RA Fleischer W., Gaasterland T., Gissi C., King B., Kochiwa H.,
RA Kuehl P., Lewis S., Matsuo Y., Nkaido I., Pesole G., Quackenbush J.,
RA Schriml L.M., Staubli F., Suzuki R., Tomita M., Wagner L., Washio T.,
RA Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barsh G.,
RA Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaldo M.F.,
RA Brownstein M.J., Bull C., Fletcher C., Fujita M., Gariboldi M.,
RA Gustincich S., Hall D., Hofmann M., Hume D.A., Kamiya M., Lee N.H.,
RA Lyons P., Marchionni L., Mashima J., Mazzaelli J., Mombaerts P.,
RA Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,

```

```

RA Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F.,
RA Suzuki H., Toyooka K., Wang K.H., Weitz C., Whittaker C., Wilming L.,
RA Wyshaw-Boris A., Yoshida K., Hasegawa Y., Kawaji H., Kohtsuki S.,
RA Hayashizaki Y.;
RT "Functional annotation of a full-length mouse cDNA collection.";
RL Nature 409:685-690(2001).
DR EMBL; AK007918; BAB25349.1; -.
DR HSP; P01842; 7FAB.
DR MGD; MGI:96443; Igh-1.
DR InterPro; IPR003599; Ig.
DR InterPro; IPR003597; Ig_c1.
DR InterPro; IPR003600; Ig_Like.
DR InterPro; IPR003006; IG_MHC.
DR InterPro; IPR003596; IG_V.
DR Pfam; PF00047; Ig; 4.
DR SMART; SM00409; IG; 2.
DR SMART; SM00407; Igc1; 3.
DR SMART; SM00406; Igv; 1.
DR SMART; SM00410; IG_Like; 1.
DR PROSITE; PS00290; IG_MHC; UNKNOWN_1.
SQ SEQUENCE 473 AA; 51699 MW; 9DED57A514475FBB CRC64;

Query Match 61.2%; Score 816; DB 11; Length 473;
Best Local Similarity 50.5%; Pred. No. 2.1e-24;
Matches 144; Conservative 38; Mismatches 42; Indels 61; Gaps 1;
QY 5 HT-----6
Db 189 HTFPALLOSLYLTSSVTVTSTWPSQITCNVAHPASSTKVKDKIEPRVPTQNPCPP 248
QY 7 ---CPCPAPELLGGPSVFLPPKPKDITLMISRTPEVTCVVVDVSHEDPEVKFNWYGV 63
Db 249 LKECPPCAAPDLLGGPSVFIFPKIKDVLMLISLPMVTCTVVVDSEDDPDVQISWFN 308
QY 64 EVHNATKPREQYNSTYRVSVLTVLHQLWLGKEYCKVSNKALPAPIETKISKAKG 123
Db 309 EVHTAQTHREDYNSTLRLVWSALPIQHDWMSGREFKCKVNNRRLPSPIETKISKPR 368
QY 124 PREQVYTLPPSRDELTKNOVSLTCLVKGYFSDIAVFNESGNQPNYKTTTPVLDSG 183
Db 369 VRAPQVYVLPPEAEEMTKKEFLTCMITGFLPAEIAVDWTSNGRTEQNKNTATVLDSG 428
QY 184 SFFLYSKLTVDKSRWQGNVFCSCVWHEALHNHYTQKSLSLSPGK 228
Db 429 SYFMYSKLRVQKSTWERSLGFACSVVHLEGLHNHLTKTSRSLGK 473

RESULT 9
Q99L25
ID Q99L25 PRELIMINARY; PRT; 473 AA.
AC Q99L25;
DT 01-JUN-2001 (TrEMBLrel. 17, Created)
DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE Similar to RIKEN cDNA 1810060009 gene.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RA Strausberg R.;
RL Submitted (FEB-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC003888; AAH03888.1; -.
DR HSP; P01842; 7FAB.
DR InterPro; IPR003599; Ig.
DR InterPro; IPR003597; Ig_c1.
DR InterPro; IPR003600; Ig_Like.
DR InterPro; IPR003006; IG_MHC.
DR InterPro; IPR003596; IG_V.
DR Pfam; PF00047; Ig; 3.
DR SMART; SM00409; IG; 2.
DR SMART; SM00407; Igc1; 3.

```

```
Query Match      68.5%; Score 914; DB 6; Length 337;
Best Local Similarity 56.8%; Pred. No. 6.6e-29;
Matches 163; Conservative 33; Mismatches 29; Indels 63; Gaps 2;

QY 5 HT-----
Db 51 HTFSPVLQSSGFYSLSSMVTVPASTWTSETYICNVVHAASNFVKDKRIEIPDNHQKVC 110
QY 7 ---CPPCAPPELLGGPSVFLPPPKDITLMSRTPEVTCVVVDVSHEDPEVKFNWYVDG 63
Db 111 MSKCPKCAPPELLGGPSVFLPPPKDITLMSRTPEVTCVVVDVSHEDPEVKFNWYMDG 170
QY 64 EVHNATKPREEQNSTYRVVSVLTVLHODWLNKGEYKCKVSNKALPAPIETKTSKAKG 123
Db 171 EVRTATTPREEQNSTYRVVSVLTVLHODWLNKGEYKCKVSNKALPAPIETKTSKAKG 230
QY 124 PREQVYTLPPSDELTKNQVSTCLVKGPSPDIAYEVWESNGQP--ENNYKTTPPVLD 181
Db 231 SQEPQVYVLAPHDELKSKSVTCLVKDFYPEINIEWQSGQPELETYKYSTTQAQDS 290
QY 182 DGSFELYSKLTVDKSRWQGNVFSVCSVMHEALHNHYTKQSLSPGK 228
Db 291 DGSFELYSKLTVDKSRWQGNVFSVCSVMHEALHNHYTKQSLSPGK 337

RESULT 5
Q8R3V9 PRELIMINARY; PRT; 469 AA.
ID Q8R3V9 AC Q8R3V9;
DT 01-JUN-2002 (TrEMBLrel. 21, Created)
DT 01-JUN-2002 (TrEMBLrel. 21, Last sequence update)
DE 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE Hypothetical 52.0 kDa protein.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RA Strausberg R.;
RL Submitted (FEB-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC024405; AAH24405.1; -
DR HSSP; P01842; 7FAB.
DR MGD; MGI:96446; Igh-4.
DR InterPro; IPR003597; Ig.
DR InterPro; IPR003597; Ig-cl.
DR InterPro; IPR003600; Ig-like.
DR InterPro; IPR003006; Ig_MHC.
DR InterPro; IPR003596; Ig_v.
DR Pfam; PF00047; Ig_4.
DR SMART; SM00409; IG; 2.
DR SMART; SM00407; IGcl; 2.
DR SMART; SM00406; IGV; 1.
DR SMART; SM00410; IG-like; 1.
DR PROSITE; PS00290; IG_MHC; UNKNOWN_1.
SQ SEQUENCE 469 AA; 51976 MW; 534793F155D05457 CRC64;

Query Match      62.7%; Score 837; DB 11; Length 469;
Best Local Similarity 52.3%; Pred. No. 2.8e-25;
Matches 145; Conservative 41; Mismatches 35; Indels 56; Gaps 3;

QY 5 HT-----
Db 196 HTFPAVLQSDLYTLSSSVTPSPSTWSPQVTCNVHPASSTKVDDKIVPRDCGCKPCICT 255
QY 12 APPELLGGPSVFLPPPKDITLMSRTPEVTCVVVDVSHEDPEVKFNWYVDGVEVHNATK 71
Db 256 VPEV---SSVFIPPPKPKDVLTTLPKVTCTVVVDISKDDPEVQFSWFVDVVEVHTAQK 312
QY 72 PREEQNSTYRVVSVLTVLHODWLNKGEYKCKVSNKALPAPIETKTSKAKGPREQVYT 131
Db 313 PREEQNSTYRVVSVLTVLHODWLNKGEYKCKVSNKALPAPIETKTSKAKGPREQVYT 372
QY 132 LPPSRDELTKNOVSLTCLVKGPSPDIAYEVWESNGQPENNYKTPPVLDSDGSFFLYSKL 191
Db 373 IPPPKQMAKDKVSLTCLMTDFFPEDITVWQWNGQPAENYKNTQPIMDTDSGFYISKL 432
QY 192 TVDKSRWQGNVFSVCSVMHEALHNHYTKQSLSPGK 228
Db 433 NVOKSNWEAGNTFTCSVLHGLHNHHTKSLSHSPGK 469

RESULT 6
Q99LC4 PRELIMINARY; PRT; 463 AA.
ID Q99LC4 AC Q99LC4;
DT 01-JUN-2001 (TrEMBLrel. 17, Created)
DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE Similar to RIKEN cDNA 1810060009 gene.
GN IGH-4.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RA Strausberg R.;
RL Submitted (FEB-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC003435; AAH03435.1; -
DR HSSP; P01842; 7FAB.
DR MGD; MGI:96446; Igh-4.
DR InterPro; IPR003597; Ig.
DR InterPro; IPR003597; Ig-cl.
DR InterPro; IPR003600; Ig-like.
DR InterPro; IPR003006; Ig_MHC.
DR InterPro; IPR003596; Ig_v.
DR Pfam; PF00047; Ig_4.
DR SMART; SM00409; IG; 2.
DR SMART; SM00407; IGcl; 2.
DR SMART; SM00406; IGV; 1.
DR SMART; SM00410; IG-like; 1.
DR PROSITE; PS00290; IG_MHC; UNKNOWN_1.
SQ SEQUENCE 463 AA; 51007 MW; EAA674C6BBC30783 CRC64;

Query Match      62.7%; Score 836; DB 11; Length 463;
Best Local Similarity 52.0%; Pred. No. 3e-25;
Matches 144; Conservative 42; Mismatches 35; Indels 56; Gaps 3;

QY 5 HT-----
Db 190 HTFPAVLQSDLYTLSSSVTPSPSTWSPQVTCNVHPASSTKVDDKIVPRDCGCKPCICT 249
QY 12 APPELLGGPSVFLPPPKDITLMSRTPEVTCVVVDVSHEDPEVKFNWYVDGVEVHNATK 71
Db 250 VPEV---SSVFIPPPKPKDVLTTLPKVTCTVVVDISKDDPEVQFSWFVDVVEVHTAQK 306
QY 72 PREEQNSTYRVVSVLTVLHODWLNKGEYKCKVSNKALPAPIETKTSKAKGPREQVYT 131
Db 307 PREEQNSTYRVVSVLTVLHODWLNKGEYKCKVSNKALPAPIETKTSKAKGPREQVYT 366
QY 132 LPPSRDELTKNOVSLTCLVKGPSPDIAYEVWESNGQPENNYKTPPVLDSDGSFFLYSKL 191
Db 367 IPPPKQMAKDKVSLTCLMTDFFPEDITVWQWNGQPAENYKNTQPIMDTDSGFYISKL 426
QY 192 TVDKSRWQGNVFSVCSVMHEALHNHYTKQSLSPGK 228
Db 427 NVOKSNWEAGNTFTCSVLHGLHNHHTKSLSHSPGK 463

RESULT 7
Q99IA4 PRELIMINARY; PRT; 437 AA.
ID Q99IA4 AC Q99IA4;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE Gammal heavy chain of Mab7 (Fragment).
GN IGH-4.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RA Wilde K.G., Yu X., Ekramoddoullah A.K.M., Misra S.;
RT "Cloning of cDNAs encoding for anti-white pine blister rust monoclonal
```

```
Db 425 DGSFFLYSKLTVDKSRWQGNVFCSCVMHEALHNHYTKQSLSPGK 471
RESULT 2
Q96PQ8 PRELIMINARY; PRT; 701 AA.
AC Q96PQ8;
DT 01-DEC-2001 (TREMBlrel. 19, Created)
DT 01-DEC-2001 (TREMBlrel. 19, Last sequence update)
DT 01-MAR-2002 (TREMBlrel. 20, Last annotation update)
DE Factor VII active site mutant immunoconjugate.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=21477448; PubMed=11593034;
RA Hu Z., Garen A.;
RT "Targeting tissue factor on tumor vascular endothelial cells and tumor
RT cells for immunotherapy in mouse models of prostatic cancer.";
RL Proc. Natl. Acad. Sci. U.S.A. 98:12180-12185(2001).
DR EMBL: AE727274; AAK58686.1; -.
DR InterPro: IPR000152; ASX_Hydroxyl.
DR InterPro: IPR000561; EGF-like.
DR InterPro: IPR000742; EGF_2.
DR InterPro: IPR001881; EGF_Ca.
DR InterPro: IPR003006; Ig_MHC.
DR InterPro: IPR001254; Ser_protease_Try.
DR InterPro: IPR000294; VitK_dep_GLA.
DR Pfam: PF00008; EGF; 2.
DR Pfam: PF00594; gla; 1.
DR Pfam: PF00047; Ig; 2.
DR Pfam: PF00089; trypsin; 1.
DR SMART: SM00181; EGF; 2.
DR PROSITE: PS00010; ASX_HYDROXYL; UNKNOWN_1.
DR PROSITE: PS00022; EGF_1; UNKNOWN_1.
DR PROSITE: PS01186; EGF_2; UNKNOWN_1.
DR PROSITE: PS01187; EGF_CA; UNKNOWN_1.
DR PROSITE: PS00011; GLU_CARBOXYLATION; UNKNOWN_1.
DR PROSITE: PS00290; IG_MHC; UNKNOWN_1.
DR PROSITE: PS00240; TRYPSIN_DOM; 1.
DR PROSITE: PS00134; TRYPSIN_HIS; UNKNOWN_1.
DR PROSITE: PS00135; TRYPSIN_SER; UNKNOWN_1.
KW Hydrolase; Serine protease.
SQ SEQUENCE 701 AA; 77826 MW; 94AC6CEB42CC992F CRC64;

Query Match 92.4%; Score 1233; DB 4; Length 701;
Best Local Similarity 100.0%; Pred. No. 5.7e-41;
Matches 227; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 DKTHTCPCPAPPELLGGPSVFLFPPPKDTLMISRTPEVTCVVVDVSHEDPEVKFNWYVD 61
Db 475 DKTHTCPCPAPPELLGGPSVFLFPPPKDTLMISRTPEVTCVVVDVSHEDPEVKFNWYVD 534
QY 62 GVEVHNAKTPREEQYNSTYRVVSVLTVLHQDLNGLKEYCKVSKNALKALPAPIEKTISKAK 121
Db 535 GVEVHNAKTPREEQYNSTYRVVSVLTVLHQDLNGLKEYCKVSKNALKALPAPIEKTISKAK 594
QY 122 GQPREPQVYTLPPSDELTKNQVSLTCLVKGFYPSDIAVEWESNGQPENNYKTPPPVLDLS 181
Db 595 GQPREPQVYTLPPSDELTKNQVSLTCLVKGFYPSDIAVEWESNGQPENNYKTPPPVLDLS 654
QY 182 DGSFFLYSKLTVDKSRWQGNVFCSCVMHEALHNHYTKQSLSPGK 228
Db 655 DGSFFLYSKLTVDKSRWQGNVFCSCVMHEALHNHYTKQSLSPGK 701

RESULT 3
Q8TC63 PRELIMINARY; PRT; 473 AA.
AC Q8TC63;
DT 01-JUN-2002 (TREMBlrel. 21, Created)

01-JUN-2002 (TREMBlrel. 21, Last sequence update)
01-JUN-2002 (TREMBlrel. 21, Last annotation update)
DE Hypothetical 52.0 kDa protein.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=98383416; PubMed=9717671;
RA Wagner B., Overesch G., Sheoran A., Holmes M., Richards C.,
RA Leibold W., Radbruch A.;
RT "Organization of the equine immunoglobulin heavy chain constant region
RT genes. III. Alignment of c-mu, c-gamma, c-epsilon and c-alpha genes.";
RL Immunobiology 199:105-119(1998).
DR EMBL: AJ300875; CAC44624.1; -.
DR InterPro: IPR003006; Ig_MHC.
DR Pfam: PF00047; Ig; 2.
DR PROSITE: PS00290; IG_MHC; UNKNOWN_2.
FT NON_TER 1
SQ SEQUENCE 337 AA; 37438 MW; A60BF2B01DEFD1F6 CRC64;
```

GenCore version 5.1.4.p5.4578  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: April 21, 2003, 10:42:21 ; Search time 36.1765 Seconds  
(without alignments)  
1401.120 Million cell updates/sec

Title: 2LINK7  
Perfect score: 1334  
Sequence: 1 MDKTHCTPCPCAPPELLGGPS.....GKDWLKAFAVDKVAEKLKEAF 246

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.0

Searched: 671580 seqs, 206047115 residues  
Total number of hits satisfying chosen parameters: 671580

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000  
Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : SPTREMBL\_21.\*

- 1: sp\_archaea.\*
- 2: sp\_bacteria.\*
- 3: sp\_fungi.\*
- 4: sp\_human.\*
- 5: sp\_invertebrate.\*
- 6: sp\_mammal.\*
- 7: sp\_mhc.\*
- 8: sp\_organelle.\*
- 9: sp\_phase.\*
- 10: sp\_plant.\*
- 11: sp\_rodent.\*
- 12: sp\_virus.\*
- 13: sp\_vertebrate.\*
- 14: sp\_unclassified.\*
- 15: sp\_rvirus.\*
- 16: sp\_bacteriap.\*
- 17: sp\_archaeap.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1233	92.4	471	4	Q8TC77
2	1233	92.4	701	4	Q96P08
3	1143	85.7	473	4	Q8TC63
4	914	68.5	337	6	Q95M34
5	837	62.7	469	11	Q8R3V9
6	836	62.7	463	11	Q99LC4
7	831	62.3	437	11	Q9R1A4
8	816	61.2	473	11	Q9D8L4
9	811	60.8	473	11	Q99L25
10	810	60.7	468	11	Q99L31
11	788	59.1	474	11	Q8R3H6
12	787	59.0	473	11	Q91Z05
13	486	36.4	34350	4	Q8WZ42
14	466	34.9	26926	4	Q10466
15	466	34.9	26926	4	Q8WZB3
16	417	31.3	17352	5	Q95YM2

17	405	30.4	16215	5	Q9NFS3
18	388	29.1	597	4	Q9BQB8
19	388	29.1	597	4	Q9BUI0
20	382	28.6	597	4	Q96BB9
21	382	28.6	618	4	Q96AA6
22	379	28.4	588	4	Q8WDX4
23	378	28.3	7962	4	Q10465
24	377	28.3	614	4	Q96GA6
25	376	28.2	613	4	Q96EV0
26	376	28.2	613	4	Q8WUK1
27	372	27.9	375	4	Q9BSZ1
28	364	27.3	613	11	Q8VCX7
29	363	27.2	6831	5	Q23550
30	363	27.2	7160	5	Q23551
31	362	27.1	15281	3	Q09164
32	356	26.7	4824	5	Q95YM1
33	356	26.7	7107	5	Q9V4F7
34	352	26.5	5636	4	Q96RW7
35	352	26.4	6658	5	Q76281
36	349	26.2	4796	5	Q9NL88
37	348	26.1	6632	5	Q01761
38	348	26.1	8563	2	Q54297
39	348	26.1	13055	5	Q09165
40	347	26.0	6632	5	Q17362
41	346	25.9	4796	5	Q9W055
42	346	25.9	5604	4	Q8WZ53
43	344	25.8	5198	5	Q76518
44	338	25.3	4816	5	Q8T103
45	337	25.3	9376	2	O85168

ALIGNMENTS

RESULT 1

Q8TC77	ID	Q8TC77	PRELIMINARY;	PRT;	471 AA.
AC	Q8TC77				
DT	01-JUN-2002	(Tremblrel. 21, Created)			
DT	01-JUN-2002	(Tremblrel. 21, Last sequence update)			
DT	01-JUN-2002	(Tremblrel. 21, Last annotation update)			
DE	Hypothetical 51.8 kDa protein.				
OS	Homo sapiens (Human).				
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;				
OC	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.				
OX	NCBI_TaxID=9606;				
RN	[1]				
RP	SEQUENCE FROM N.A.				
RC	TISSUE=SPLEEN;				
RA	Strausberg R;				
RL	Submitted (FEB-2002) to the EMBL/GenBank/DBJ databases.				
DR	EMBL; BC024289; AAH24289.1; -				
KW	Hypothetical protein.				
SQ	SEQUENCE 471 AA; 51791 MW; 388F74CF588660E CRC64;				

Query Match	92.4%;	Score 1233;	DB 4;	Length 471;
Best Local Similarity	100.0%;	Pred. No. 1.5e-41;		
Matches 227;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;
QY	2	DKTHCTPCPCAPPELLGGPSVFLFPPKPKDTLMISRTPEVTCVVVDVSHEDPEVKFNMYVD	61	
Db	245	DKTHCTPCPCAPPELLGGPSVFLFPPKPKDTLMISRTPEVTCVVVDVSHEDPEVKFNMYVD	304	
QY	62	GVEVHNKAKTPREEQYNSTYRVVSVLTVLHQDWLNGKEYKCKVSNKALPAPIEKTISKAK	121	
Db	305	GVEVHNKAKTPREEQYNSTYRVVSVLTVLHQDWLNGKEYKCKVSNKALPAPIEKTISKAK	364	
QY	122	GQREPOVYTLPPSRDELTKNOVSLTCLVKGFYPSDIAVWESNGQPENNYKTPPPVLD	181	
Db	365	GQREPOVYTLPPSRDELTKNOVSLTCLVKGFYPSDIAVWESNGQPENNYKTPPPVLD	424	
QY	182	DGSFFLYSKLTVDKSRWQQGNVSCFVSHEALHNHYTQKSLSLSPGK	228	

```
FT DISULFID 110 110 INTERCHAIN (WITH A HEAVY CHAIN).
FT DISULFID 112 112 INTERCHAIN (WITH A HEAVY CHAIN).
FT DISULFID 144 204
FT DISULFID 250 308
FT MOD.RES 330 330 REMOVED POST-TRANSLATIONALLY.
SQ SEQUENCE 330 AA; 36389 MW; B84361C5445A6864 CRC64;

Query Match 60.9%; Score 813; DB 1; Length 330;
Best Local Similarity 52.9%; Pred. No. 2.4e-20;
Matches 148; Conservative 30; Mismatches 46; Indels 56; Gaps 2;
Qy 5 HT-----
Db 51 HTFPAVLQSDLYTLSSSVTVTSSTWPSQSIITCNVAHPASSTKVDKKIEPRGPTIKPCPPC 110
Qy 10 -CPAPPELLGGPSVFLFPPKPKDTLMISRTPEVTCVVVDVSHEDPEVKFNWYVDGVEVHNA 68
Db 111 KCPAPNLLGGPSVFIFPPKIKDVLMLISLPIVTCVVVDVSEDDPDVQISWFWNNVEVHTA 170
Qy 69 KTKPREOYNSTYRVVSVLTVLHQDWLNGKEYKCKVSNKALPAPIEKTISKAKGQPREPQ 128
Db 171 QTQTHREDYNSTLRVVSALPIQHODWMSGKEFKCKVNNKDLPAPIERTISKPKGSVRAPQ 230
Qy 129 VYTLPPSRDELTKNOVSLTCLVKGYFSPDIAVEWESNQPENNYKTTPVLDSDGSFELY 188
Db 231 VVVLPPPEEEMTKQVTLTCMTDFMPEDIYVEWTNNGKTELNYKNTPEVLDSDGSYFMY 290
Qy 189 SKLTVDKSRWQGNVFCSSVMHEALHNHYTOKSLSLSPGK 228
Db 291 SKLRVEKKNWVERNSYCSVVHIEGLHNHHTTKSFSRTPEGK 330

Search completed: April 21, 2003, 10:43:10
Job time : 10.4059 secs
```

DT 01-FEB-1991 (Rel. 17, Last sequence update)  
 DT 15-JUL-1999 (Rel. 38, Last annotation update)  
 DE Ig gamma-2C chain C region.  
 OS Rattus norvegicus (Rat).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.  
 OX NCBI\_TaxID=10116;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=88166903; PubMed=3127222;  
 RA Brueggemann M., Delmastro-Galfré P., Waldmann H., Calabi F.;  
 RT "Sequence of a rat immunoglobulin gamma 2c heavy chain constant  
 region cDNA: extensive homology to mouse gamma 3.";  
 RL Eur. J. Immunol. 18:317-319(1988).  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announcement/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 CC EMBL; X07189; CAA30169.1; -  
 DR PIR; S00847; S00847.  
 DR HSSP; P01842; 7FAB.  
 DR InterPro; IPR003006; Ig\_MHC.  
 DR InterPro; IPR003597; Ig-cl.  
 DR InterPro; IPR003600; Ig-like.  
 DR Pfam; PF00047; Ig; 2.  
 DR SMART; SM00410; IG-like; 1.  
 DR SMART; SM00407; IGcl; 2.  
 DR PROSITE; PS00290; IG\_MHC; 1.  
 DR Immunoglobulin domain; Immunoglobulin C region.  
 FT NON\_TER 1 1  
 FT DOMAIN 1 97 CH1.  
 FT DOMAIN 98 113 HINGE.  
 FT DOMAIN 114 222 CH2.  
 FT DOMAIN 223 329 CH3.  
 FT DISULFID 15 15 INTERCHAIN (WITH A LIGHT CHAIN).  
 FT DISULFID 27 82  
 FT DISULFID 111 111 INTERCHAIN (WITH A HEAVY CHAIN).  
 FT DISULFID 113 113 INTERCHAIN (WITH A HEAVY CHAIN).  
 FT DISULFID 143 203  
 FT DISULFID 249 307  
 SQ SEQUENCE 329 AA; 36571 MW; 5FCD7B7933850773 CRC64;  
 Query Match 61.3%; Score 818; DB 1; Length 329;  
 Best Local Similarity 57.9%; Pred. No. 1.6e-20;  
 Matches 146; Conservative 42; Mismatches 38; Indels 26; Gaps 2;  
 QY 3 KTHTC-----PP---CPAPELLGGPSVFLFPPKPKDTLMISR 36  
 DB QTVCVSVAHPATKSNLIKRIEPRRPKRPTDTCSDDLGRSVFIFFPKPKDILMITL 137  
 QY 37 TPETVCVVVDVSHEDPEVFANFVVDGVVEVHNAKTKPREEQYNSTYRVSVLTVHLQDWLN 96  
 DB 138 TPRTVCVVVDVSEEDPQSFNEVDNVRVFTAGTQPEHEQLNGTFRVSTLHIOHODWMS 197  
 QY 97 GREYKCKVSKALPAPTEIKTSKAKQPREPQVYTLPPSRDELTKNQVSLTCLVKGFYPS 156  
 DB 198 GKEFKCKVANKDLPSPIETKTSKPRGKARTPQVYTIPTPPREQMKNKNSLTCTVTSFYPA 257  
 QY 157 DIAVENESGOPENNKKTPPVLDSDGSFPLYSLKLTVDKSRVQGNVFCFSVWHEALHNN 216  
 DB 258 SISVERNGEGLFQDYKTLPLVLDSDSESYFLYSKLSVDTSWMRGDIYTCVVYHEALHNN 317  
 QY 217 YTKSLSLSPGK 228  
 DB 318 HTQKNLSRSPGK 329  
 RESULT 15

GCAN\_MOUSE  
 ID GCAA\_MOUSE STANDARD; PRT; 330 AA.  
 AC P01863;  
 DT 21-JUL-1986 (Rel. 01, Created)  
 DT 21-JUL-1986 (Rel. 01, Last sequence update)  
 DT 30-MAY-2000 (Rel. 39, Last annotation update)  
 DE Ig gamma-2A chain C region, A allele.  
 DE Mus musculus (Mouse)  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 OX NCBI\_TaxID=10090;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=81076554; PubMed=6777755;  
 RA Sikorav J.-L., Auffray C., Rougeon F.;  
 RT "Structure of the constant and 3' untranslated regions of the murine  
 RT Balb/c gamma 2a heavy chain messenger RNA.";  
 RL Nucleic Acids Res. 8:3143-3155(1980).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=81198976; PubMed=6262729;  
 RA Yamawaki-Kataoka Y., Miyata T., Honjo T.;  
 RT "The complete nucleotide sequence of mouse immunoglobulin gamma 2a gene  
 RT and evolution of heavy chain genes: further evidence for intervening  
 RT sequence-mediated domain transfer.";  
 RL Nucleic Acids Res. 9:1365-1381(1981).  
 RN [3]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=81223894; PubMed=6787604;  
 RA Ollo R., Auffray C., Morchamps C., Rougeon F.;  
 RT "Comparison of mouse immunoglobulin gamma 2a and gamma 2b chain genes  
 RT suggests that exons can be exchanged between genes in a multigenic  
 RT family.";  
 RL Proc. Natl. Acad. Sci. U.S.A. 78:2442-2446(1981).  
 RN [4]  
 RP MYELOMA PROTEIN MOPC 173.  
 RX MEDLINE=74175517; PubMed=4831970;  
 RA Bourgois A., Fougereau M., Rocca-Serra J.;  
 RT "Determination of the primary structure of a mouse IgG2a  
 RT immunoglobulin: amino-acid sequence of the Fc fragment. Implications  
 RT for the evolution of immunoglobulin structure and function.";  
 RL Eur. J. Biochem. 43:423-435(1974).  
 RN [5]  
 RP DISULFIDE BONDS.  
 RX MEDLINE=73056887; PubMed=4565406;  
 RA de Preval C., Fougereau M.;  
 RT "Determination of the primary structure of a mouse gamma G2a  
 RT immunoglobulin. Identification of the disulfide bridges.";  
 RL Eur. J. Biochem. 30:452-462(1972).  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announcement/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 CC EMBL; V00798; CAA24178.1; -  
 DR PIR; A02152; G2MSA.  
 DR HSSP; P01842; 7FAB.  
 DR InterPro; IPR003006; Ig\_MHC.  
 DR InterPro; IPR003597; Ig-cl.  
 DR InterPro; IPR003600; Ig-like.  
 DR Pfam; PF00047; Ig; 2.  
 DR SMART; SM00410; IG-like; 1.  
 DR SMART; SM00407; IGcl; 2.  
 DR PROSITE; PS00290; IG\_MHC; 1.  
 DR Immunoglobulin domain; Immunoglobulin C region.  
 FT NON\_TER 1 1  
 FT DISULFID 15 15 INTERCHAIN (WITH A LIGHT CHAIN).  
 FT DISULFID 27 82  
 FT DISULFID 107 107 INTERCHAIN (WITH A HEAVY CHAIN).  
 FT DISULFID 107 107

```
DR EMBL; V00793; CAA24174.1; .
DR PIR; B02159; G1NSM.
DR HSP; P01842; 7FAB.
DR MGD; MGI:96446; Igh-4.
DR InterPro; IPR003006; Ig_MHC.
DR InterPro; IPR003597; Ig_c1.
DR Pfam; PF00047; Ig; 3.
DR SMART; SM00407; Igc1; 2.
DR PROSITE; PS00290; IG_MHC; 1.
KW Immunoglobulin domain; Immunoglobulin C region; Glycoprotein;
KW Alternative splicing; Transmembrane.
FT NON_TER 1
FT DOMAIN 1 97 CH1.
FT DOMAIN 98 110 HINGE.
FT DOMAIN 111 217 CH2.
FT DOMAIN 218 324 CH3.
FT DISULFID 27 82
FT DISULFID 102 102 INTERCHAIN (WITH A LIGHT CHAIN).
FT DISULFID 104 104 INTERCHAIN (WITH A HEAVY CHAIN).
FT DISULFID 107 107 INTERCHAIN (WITH A HEAVY CHAIN).
FT DISULFID 109 109 INTERCHAIN (WITH A HEAVY CHAIN).
FT DISULFID 138 198
FT CARBOHYD 174 174 N-LINKED (GLCNAC. . .).
FT DISULFID 244 302
FT TRANSMEM 340 357 POTENTIAL.
FT DOMAIN 338 393 CYTOPLASMIC (POTENTIAL).
SQ SEQUENCE 393 AA; 43386 MW; 4CC88343B7A1CE27 CRC64;

Query Match 61.7%; Score 823; DB 1; Length 393;
Best Local Similarity 51.8%; Pred. No. 1.9e-20;
Matches 145; Conservative 47; Mismatches 42; Indels 46; Gaps 5;

QY 7 CPP--CPAPELLGGPSVFLFPKPKDLMISRTPEVTCVVVDVSHEDPEVKENWYDGYE 64
DB 104 CKPCICTVEV---SSVFIEFPKPKDVLITLTPKVTVCVVVDISKDDPEVQSFWEVDGYE 160
QY 65 VHNATKPREEQNSTYRVSVLTVLHQDLNKGKEYCKVSNKALPAPIERTISKAKGAP 124
DB 161 VHTAQTPREEQNSTFRSVSELPIMHQDLNKGKEYCKVSNKALPAPIERTISKAKGAP 220
QY 125 REPQVYTLPPSRDELTKNOVSLTCLVKGPSPDIAVEWESNGQPNENYKTTTPVLDSDGS 184
DB 221 KAPQVYTIPTPEQNAKDKVSLTCHITDFPEIDIVWQWNGQPAENYKNTQPIWNTGS 280
QY 185 FFYLSKLTVDKSRWQGNVFCSSVMHEALHNHYTKSLSPGKD----- 229
DB 281 YFYSLKLVOKSNWEAGNTFTCSVLHLEGLHNHTKSLSPGLQDTCACAAQDGLDG 340
QY 230 -W-----LKAFYDKVAEKLKE 244
DB 341 LMTTITIFISLLSVCSAAVTLFKVKWIFSSVVE-LKQ 379

RESULT 13
GCAM_MOUSE
ID GCAM_MOUSE STANDARD; PRT; 399 AA.
AC P01865;
DT 21-JUL-1986 (Rel. 01, Created)
DT 01-AUG-1991 (Rel. 19, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE Ig gamma-2a chain C region, membrane-bound form.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=82222190; PubMed=6283537;
RA Yamawaki-Kataoka Y., Nakai S., Miyata T., Honjo T.;
RT "Nucleotide sequences of gene segments encoding membrane domains of
RT immunoglobulin gamma chains."
RL Proc. Natl. Acad. Sci. U.S.A. 79:2623-2627(1982).
CC -|- ALTERNATIVE PRODUCTS: CELL LINES PRODUCING IGG CONTAIN TWO MRNA
```

```
CC SPECIES FOR IG GAMMA CHAINS. THE MAJOR SPECIES ENCODES SECRETED
CC GAMMA CHAINS. A LESS ABUNDANT SPECIES APPEARS TO ENCODE MEMBRANE-
CC BOUND CHAINS IN THAT IT CONTAINS AN ALTERNATIVE 3' END, ENCODED
CC IN SEPARATE EXONS, THAT IS HOMOLOGOUS WITH THE MEMBRANE-BOUND
CC SEGMENT OF MU CHAINS.
CC -|- MISCELLANEOUS: THE SEQUENCE OF RESIDUES 1-329 IS ASSUMED TO BE
CC IDENTICAL WITH THE CORRESPONDING REGION OF THE SECRETED FORM OF
CC THE A ALLELE.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; J00471; AAB59661.1; ALT_INIT.
CC PIR; A02154; G2MSAM.
CC HSP; P01842; 7FAB.
CC MGD; MGI:96443; Igh-1.
CC InterPro; IPR003006; Ig_MHC.
CC InterPro; IPR003597; Ig_c1.
CC InterPro; IPR003600; Ig_Like.
CC Pfam; PF00047; Ig; 2.
CC SMART; SM00410; Igc1; 1.
CC SMART; SM00407; Igc1; 2.
CC PROSITE; PS00290; IG_MHC; 1.
KW Immunoglobulin domain; Immunoglobulin C region; Glycoprotein;
KW Transmembrane; Alternative splicing.
FT NON_TER 1
FT DISULFID 15 15 INTERCHAIN (WITH A LIGHT CHAIN).
FT DISULFID 27 82 INTERCHAIN (WITH A HEAVY CHAIN).
FT DISULFID 107 107 INTERCHAIN (WITH A HEAVY CHAIN).
FT DISULFID 110 110 INTERCHAIN (WITH A HEAVY CHAIN).
FT DISULFID 112 112 INTERCHAIN (WITH A HEAVY CHAIN).
FT DISULFID 144 204
FT DISULFID 250 308
FT TRANSMEM 346 363 POTENTIAL.
FT DOMAIN 364 399 CYTOPLASMIC (POTENTIAL).
FT CARBOHYD 180 180 N-LINKED (GLCNAC. .) (POTENTIAL).
SQ SEQUENCE 399 AA; 44020 MW; 4C38138FAED3FF0 CRC64;

Query Match 61.4%; Score 819; DB 1; Length 399;
Best Local Similarity 54.3%; Pred. No. 2.7e-20;
Matches 152; Conservative 33; Mismatches 52; Indels 43; Gaps 4;

QY 7 CPP--CPAPELLGGPSVFLFPKPKDLMISRTPEVTCVVVDVSHEDPEVKENWYDGYE 64
DB 107 CPPCKPAPNLLGGPSVFIFFPKPKIDVLMISLSPITVCVVVDVSDDDPQISWFVNNYE 166
QY 65 VHNATKPREEQNSTYRVSVLTVLHQDLNKGKEYCKVSNKALPAPIERTISKAKGAP 124
DB 167 VHTAQTPREEQNSTLRVVSALPIQHDMSGKECKVSNKALPAPIERTISKPKGV 226
QY 125 REPQVYTLPPSRDELTKNOVSLTCLVKGPSPDIAVEWESNGQPNENYKTTTPVLDSDGS 184
DB 227 RAPQVYVLPPEEEMTKQVTLTCMVTDPMPEIDIVWQWNGKTELNKNTEPVLSDGS 286
QY 185 FFYLSKLTVDKSRWQGNVFCSSVMHEALHNHYTKSLSPGKD----- 229
DB 287 YFYSLKLVOKSNWEAGNTFTCSVLHLEGLHNHTKSLSPGLDLDVCAEAQDGLDG 346
QY 230 -W-----LKAFYDKVAEKLKE 244
DB 347 LMTTITIFISLLSVCSAAVTLFKVKWIFSSVVE-LKQ 385

RESULT 14
GCC_RAT
ID GCC_RAT STANDARD; PRT; 329 AA.
AC P20762;
DT 01-FEB-1991 (Rel. 17, Created)
```





DR EMBL; V01526; CAA24767.1; ALT\_SEQ.  
 DR PIR; A02155; G3MSM.  
 DR HSSP; P01857; IFC1.  
 DR InterPro; IPR003006; Ig\_MHC.  
 DR InterPro; IPR003597; Ig\_C1.  
 DR InterPro; IPR003600; Ig\_Like.  
 DR Pfam; PF00047; Ig; 3.  
 DR SMART; SM00410; Ig\_Like; 1.  
 DR SMART; SM00407; IGc1; 2.  
 DR PROSITE; PS00290; IG\_MHC; 1.  
 KW Immunoglobulin domain; Immunoglobulin C region; Glycoprotein;  
 Transmembrane; Alternative splicing.  
 FT NON\_TER 1 1  
 FT DOMAIN 1 97 CH1.  
 FT DOMAIN 98 113 HINGE.  
 FT DOMAIN 114 223 CH2.  
 FT DOMAIN 224 327 CH3.  
 FT TRANSMEM 346 362 POTENTIAL.  
 FT DOMAIN 363 398 CYTOPLASMIC (POTENTIAL).  
 FT CONFLICT 333 333 E -> G (IN REF. 2).  
 FT CONFLICT 342 342 E -> Q (IN REF. 2).  
 FT CONFLICT 388 388 P -> F (IN REF. 2).  
 SQ SEQUENCE 398 AA; 43929 MW; CF7F264B50A41B95 CRC64;

Query Match 62.8%; Score 838; DB 1; Length 398;  
 Best Local Similarity 50.5%; Pred. No. 6e-21;  
 Matches 151; Conservative 37; Mismatches 41; Indels 70; Gaps 3;

QY 3 KTHCTC-----PP---CPAPPELLGGSPVFLPPKPKDTLMIS 35  
 DB 77 QTVCINVAHPASKTELKRIEPRIPKPTPGSSCPPNIIIGSPVFFPPKPKDLMIS 136  
 QY 36 RPEVTCVVDVSHEDPEVKVNWYDGVVHNAKTKPREEQYNSTYRVSVLTVLHQQDWL 95  
 DB 137 LTPKVTVCVVDVSEDDPDVHVSFWFDNKEVHTAWTPQPREAQYNSTFRVVSALPIHQDWM 196  
 QY 96 NGKEYCKVKSNKALPAPIETKISAKAGQPREPOVYTLPPSRDELTKNOVSLTCLVKGY 155  
 DB 197 RGKEPKCKVNNKALPAPIETKISAKAGQPREPOVYTLPPSRDELTKNOVSLTCLVKGY 256  
 QY 156 SDIAVWESNGOPENNYKTPPVLDSGSEFFLYSKLTVDKSRWQGNVFCSCVMHEALHN 215  
 DB 257 EASVWERNGELEQDYKNTPPILDSGDTYFLYSLKLTVDTSWLGGEFTCSVHEALHN 316  
 QY 216 HYTKSLSLSPKGD-----WL 231  
 DB 317 HHTQNLRSRSPLELNETCAEAQDGLDGLWTITIFISLFLSVCYASVTLFKVKWI 375

RESULT 10  
 GCL\_RAT  
 ID GCL\_RAT STANDARD; PRT; 326 AA.  
 AC P20759;  
 DT 01-FEB-1991 (Rel. 17, Created)  
 DT 01-FEB-1991 (Rel. 17, Last sequence update)  
 DT 15-JUL-1999 (Rel. 38, Last annotation update)  
 DE Ig gamma-1 chain C region.  
 OS Rattus norvegicus (Rat).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.  
 OX NCBI\_TaxID=10116;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=89232738; PubMed=3149946;  
 RA Brueggemann M.;  
 RT "Evolution of the rat immunoglobulin gamma heavy-chain gene family."  
 RL Gene 74:473-482(1989).  
 DR PIR; PS0017; PS0017.  
 DR HSSP; P01842; 7FAB.  
 DR InterPro; IPR003006; Ig\_MHC.  
 DR Pfam; PF00047; Ig; 3.  
 DR SMART; SM00407; IGc1; 2.

DR PROSITE; PS00290; IG\_MHC; 1.  
 KW Immunoglobulin domain; Immunoglobulin C region; Glycoprotein.  
 FT NON\_TER 1 1  
 FT DOMAIN 1 97 CH1.  
 FT DOMAIN 98 112 HINGE.  
 FT DOMAIN 113 219 CH2.  
 FT DOMAIN 220 326 CH3.  
 FT DISULFID 27 82  
 FT DISULFID 102 102 INTERCHAIN (WITH A HEAVY CHAIN).  
 FT DISULFID 106 106 INTERCHAIN (WITH A HEAVY CHAIN).  
 FT DISULFID 109 109 INTERCHAIN (WITH A HEAVY CHAIN).  
 FT DISULFID 111 111 INTERCHAIN (WITH A HEAVY CHAIN).  
 FT DISULFID 140 200  
 FT DISULFID 246 304  
 FT CARBOHYD 176 176  
 SQ SEQUENCE 326 AA; 35946 MW; 013BAB45EF49B9DA CRC64;

Query Match 61.8%; Score 825; DB 1; Length 326;  
 Best Local Similarity 51.4%; Pred. No. 9.1e-21;  
 Matches 144; Conservative 41; Mismatches 35; Indels 60; Gaps 3;

QY 5 HT-----CPGPCP 11  
 DB 51 HTFPAVLQSLYLTSSVTPSSWPSQTVTCNVAHPASSTKVDDKIYPRNCGGDKPC- 109  
 QY 12 APELLGG---PSVFLFPKPKDITLMISRTPEVTCVVVDVSHEDPEVKFNWYDGVVEHNA 68  
 DB 110 ---ICTGSEVSVFFPPKPKDVLITLTPKVTVCVVDVSHEDPEVKFNWYDGVVEHNA 166  
 QY 69 KTKPREEQYNSTYRVSVLTVLHQQDWLNGKEYCKVKSNKALPAPIETKISAKAGQPREPO 128  
 DB 167 QTRPPEQFNSTFRVSVSELPILHQQDWLNGRTFRKVTSAAPSPLEKTSKPEGTQVPH 226  
 QY 129 VYTLPPSRDELTKNOVSLTCLVKGYPSDIAVWESNGOPENNYKTPPVLDSGSEFFLY 188  
 DB 227 VYTSPTKEMTQNEYSITCMVKGYPPDIYVWQGNQGNQENYKNTPTDGTGSEFFLY 286  
 QY 189 SKLTVDKSRWQGNVFCSCVMHEALHNHYTKSLSLSPK 228  
 DB 287 SKLVNKKKQWQGNFTCSVLHGLEHNNHTEKLSHSPGK 326

RESULT 11  
 GCL\_MOUSE  
 ID GCL\_MOUSE STANDARD; PRT; 324 AA.  
 AC P01868;  
 DT 21-JUL-1986 (Rel. 01, Created)  
 DT 21-JUL-1986 (Rel. 01, Last sequence update)  
 DT 30-MAY-2000 (Rel. 39, Last annotation update)  
 DE Ig gamma-1 chain C region.  
 OS Mus musculus (Mouse).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 OX NCBI\_TaxID=10090;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=80045036; PubMed=115593;  
 RA Honjo T., Obata M., Yamawaki-Kataoka Y., Kataoka T., Kawakami T.,  
 RA Takahashi N., Mano Y.;  
 RT "Cloning and complete nucleotide sequence of mouse immunoglobulin  
 gamma 1 chain gene."  
 RL Cell 18:559-568(1979).  
 RN [2]  
 RP SEQUENCE OF 76-324 FROM N.A. (MYELOMA PROTEIN MOPC 31C).  
 RX MEDLINE=80202559; PubMed=6769752;  
 RA Obata M., Yamawaki-Kataoka Y., Takahashi N., Kataoka T., Shimizu A.,  
 RA Mano Y., Seidman J.G., Peterlin B.M., Leder P., Honjo T.;  
 RT "Immunoglobulin gamma 1 heavy chain gene: structural gene sequences  
 cloned in a bacterial plasmid."  
 RL Gene 9:87-97(1980).  
 RN [3]  
 RP SEQUENCE OF 70-322 FROM N.A. (MYELOMA PROTEIN MOPC 21).  
 RX MEDLINE=80012837; PubMed=113776;

```
CC ----- EMBL; J00451; -; NOT_ANNOTATED_CDS.
DR PIR; B02156; G3MSC.
DR HSSP; P01857; 1FC1.
DR InterPro; IPR003006; Ig_MHC.
DR InterPro; IPR003597; Ig_c1.
DR InterPro; IPR003600; Ig_like.
DR Pfam; PF00047; Ig; 3.
DR SMART; SM00410; Ig_like; 1.
DR SMART; SM00407; Ig_c1; 2.
DR PROSITE; PS00290; IG_MHC; 1.
KW Immunoglobulin domain; Immunoglobulin C region; Glycoprotein;
KW Transmembrane; Alternative splicing.
FT NON_TER 1 1
FT DOMAIN 1 97 CH1.
FT DOMAIN 98 113 HINGE.
FT DOMAIN 114 223 CH2.
FT DOMAIN 224 327 CH3.
SQ SEQUENCE 329 AA; 36228 MW; F45827174182BAD6 CRC64;

Query Match 63.6%; Score 848; DB 1; Length 329;
Best Local Similarity 60.1%; Pred. No. 1.5e-21;
Matches 152; Conservative 35; Mismatches 39; Indels 27; Gaps 2;

QY 3 KTHC-----pp---CPAPELLGGPSVFLPPKPKDLMIS 35
DB 77 QTICNVAHSPAKTELKRIEPIKPSPPGSCPPGNILGGPSVFIPPKPKDALMIS 136
QY 36 RTEPTCVVVDVSHEDPEVKFNMYVDGVEVHNNAKTPREEQYNSTYRVVSVLTVLHQDWL 95
DB 137 LTPKVTVCVVVDSEDDPDVHVSFEVDNKEVHTAQTQPREAQYNSTFRVVSALPIQHDWM 196
QY 96 NGKEYCKVSNKALPAPIEKTISKAKGQPREPOVYITLPPSRDELTKNOVSLTCLVKGFTP 155
DB 197 RGKEFKCKVNNKALPAPIERTISKPKRAQTQVYITPPPREQMSKKVSLTCLVTNFFS 256
QY 156 SDIAVEWESNGPENNYKTPPVLDSDGSFFLYSKLTVDKSRWQGNVFCSCVMHEALHN 215
DB 257 EALSVEWERNLEQDKYKNTPTPLDSDGTYFLYSLKLTVDTSWLGQEIETCSVVHEALHN 316
QY 216 HYTKSLSLSPGK 228
DB 317 HHTQKNLSRSPGK 329

RESULT 8
GCB_RAT STANDARD; PRT; 333 AA.
AC P20761;
DT 01-FEB-1991 (Rel. 17, Created)
DT 01-FEB-1991 (Rel. 17, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE Ig gamma-2B chain C region.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=89232738; PubMed=3149946;
RA Brueggemann M.;
RT "Evolution of the rat immunoglobulin gamma heavy-chain gene family.";
RL Gene 74:473-482(1988).
DR PIR; PS0018; PS0018.
DR HSSP; P01842; 7FAB.
DR InterPro; IPR003006; Ig_MHC.
DR InterPro; IPR003597; Ig_c1.
DR InterPro; IPR003600; Ig_like.
DR Pfam; PF00047; Ig; 3.
DR SMART; SM00410; Ig_like; 1.
DR SMART; SM00407; Ig_c1; 2.
DR PROSITE; PS00290; IG_MHC; 1.
KW Immunoglobulin domain; Immunoglobulin C region.
```

```
FT NON_TER 1 1
FT DISULFID 15 15 INTERCHAIN (WITH A LIGHT CHAIN).
FT DISULFID 27 80
FT DISULFID 106 106 INTERCHAIN (WITH A HEAVY CHAIN).
FT DISULFID 109 109 INTERCHAIN (WITH A HEAVY CHAIN).
FT DISULFID 112 112 INTERCHAIN (WITH A HEAVY CHAIN).
FT DISULFID 115 115 INTERCHAIN (WITH A HEAVY CHAIN).
FT DISULFID 147 207
FT DISULFID 253 311
SQ SEQUENCE 333 AA; 55P8B64D48D460A6 CRC64;

Query Match 63.0%; Score 840; DB 1; Length 333;
Best Local Similarity 62.1%; Pred. No. 3e-21;
Matches 151; Conservative 34; Mismatches 43; Indels 15; Gaps 2;

QY 1 MDKT-----HTCP-----PCPAPELLGGPSVFLPPKPKDLMISRTPEVTCVV 45
DB 91 VDKKVERRNGGIGHKCPCTCHCKCPVPELLGGPSVFIPPKPKDILLISQNAKVTCCVV 150
QY 46 DYSHEDPEVKFNMYVDGVEVHNNAKTPREEQYNSTYRVVSVLTVLHQDWLNGKEYCKVS 105
DB 151 DYSEEDPDVQFSMFVNNEVHTAQTQPREQYNSTFRVVSALPIQHDWMNSGKEFKCKVN 210
QY 106 NKALPAPIEKTISKAKGQPREPOVYITLPPSRDELTKNOVSLTCLVKGFTVPSDIAVEWESN 165
DB 211 NKALPSPIEKTISKPKGLVKKPOVYVNGPTEQLTQVSLTCLTSGFLPNDIGVETSN 270
QY 166 GOPENNYKTPPVLDSDGSFFLYSKLTVDKSRWQGNVFCSCVMHEALHNHYTKSLSL 225
DB 271 GHIEKYNKTEPVMDSDGSFFMYSKLNVSRWDSRAPFVCSVVHGLHNHHVKEKTSRP 330
QY 226 PGK 228
DB 331 PGK 333

RESULT 9
GCB_MOUSE STANDARD; PRT; 398 AA.
AC P03987;
DT 23-OCT-1986 (Rel. 02, Created)
DT 01-AUG-1991 (Rel. 19, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE Ig gamma-3 chain C region, membrane-bound form.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Cranata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=85027161; PubMed=6092053;
RA Wells J.A., Word C.J., Rimm D., Der-Balan G.P., Martinez H.M.,
RA Tucker P.W., Blattner F.R.;
RT "Structural analysis of the murine IgG3 constant region gene.";
RL EMBO J. 3:2041-2046(1984).
RN [2]
RP SEQUENCE OF 328-398 FROM N.A.
RX MEDLINE=84041483; PubMed=6314258;
RA Komaromy M., Clayton L., Rogers J., Robertson S., Kettman J.,
RA Wall R.;
RT "The structure of the mouse immunoglobulin in gamma 3 membrane gene
segment.";
RL Nucleic Acids Res. 11:6775-6785(1983).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DB EMBL; J00451; AAB59655.1; -.
```

SQ SEQUENCE 323 AA; 35404 MW; 69EBA118D579A8B CRC64;  
 Query Match 69.0%; Score 921; DB 1; Length 323;  
 Best Local Similarity 71.7%; Pred. No. 4.7e-24;  
 Matches 167; Conservative 29; Mismatches 32; Indels 5; Gaps 2;  
 QY 1 MDKT---HTC---PPCPAPPELLGGPSVFLFPPKPKDTLMISRTPEVTCVVDVDSHEDPEVK 55  
 DB 91 VDKTVAISTCKSPKCPPELLGGPSVFLFPPKPKDTLMISRTPEVTCVVDVDSODDEVEQ 150  
 QY 56 FNTWVDGVEVHNATKPREQYNSTYRVSVLTVLHODWLNGLKREYKCKVSNKALPAPIEK 115  
 DB 151 FTWYINNEQVTRAPPLREQGFNSTIRVSTPLTHQDWLGRKEFKCKVHNKALPAPIEK 210  
 QY 116 TISKAKGPREQVYTLPPSRDELTKNQVSLTCLVKGFYPSDIAVEWESNGQPENNYKTT 175  
 DB 211 TISKARQPLEKPYTMGPPEELSSRSVSLTCMNGFYPSDIAVEWESNGKAEADNYKTT 270  
 QY 176 PVLVDSGDFLYSKLVKRVDSKRMQGNVFCVSMHEALHNHYTKLSLSPGK 228  
 DB 271 PAVLDSGDFLYNKLKSVPTSEWQGDVFTCSVMHEALHNHYTKLSRSRSPGK 323  
 RESULT 6  
 GC2\_CAVPO STANDARD; PRT; 329 AA.  
 ID GC2\_CAVPO  
 AC P01862;  
 DT 21-JUL-1986 (Rel. 01, Created)  
 DT 21-JUL-1986 (Rel. 01, Last sequence update)  
 DT 15-JUL-1999 (Rel. 38, Last annotation update)  
 DE Ig gamma-2 chain C region.  
 OS Cavia porcellus (Guinea pig).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Myricognathi; Caviidae; Cavia.  
 OX NCBI\_TaxID=10141;  
 RN [1]  
 RP SEQUENCE OF 1-3.  
 RA Trischmann T.M.;  
 RL Submitted (APR-1975) to the PIR data bank.  
 RN [2]  
 RP SEQUENCE OF 4-68.  
 RX MEDLINE=71058471; PubMed=5538606;  
 RA Birstein B.K., Hussain Q.Z., Cebra J.J.;  
 RT "Structure of heavy chain from strain 13 guinea pig  
 immunoglobulin-G(2). 3. Amino acid sequence of the region around the  
 half-cysteine joining heavy and light chains.";  
 RL Biochemistry 10:18-25(1971).  
 RN [3]  
 RP SEQUENCE OF 69-133 AND 312-329.  
 RX MEDLINE=71058486; PubMed=5538616;  
 RA Turner K.J., Cebra J.J.;  
 RT "Structure of heavy chain from strain 13 guinea pig  
 immunoglobulin-G(2). II. Amino acid sequence of the carboxyl-terminal  
 and hinge region cyanoogen bromide fragments.";  
 RL Biochemistry 10:9-17(1971).  
 RN [4]  
 RP SEQUENCE OF 134-226.  
 RX MEDLINE=75036072; PubMed=4429665;  
 RA Tracey D.E., Cebra J.J.;  
 RT "Primary structure of the CH2 homology region from guinea pig IgG2  
 antibodies.";  
 RL Biochemistry 13:4796-4803(1974).  
 RN [5]  
 RP SEQUENCE OF 227-311.  
 RX MEDLINE=75036073; PubMed=4609467;  
 RA Trischmann T.M., Cebra J.J.;  
 RT "Primary structure of the CH3 homology region from guinea pig IgG2  
 antibodies.";  
 RL Biochemistry 13:4804-4811(1974).  
 RN [6]  
 RP DISULFIDE BONDS.  
 RX MEDLINE=71058474; PubMed=4922544;  
 RA Oliveira B., Lamm M.E.;

RT "Interchain disulfide bridges of guinea pig gamma-2-immunoglobulin.";  
 RL Biochemistry 10:26-31(1971).  
 CC -I- MISCELLANEOUS: THIS CHAIN WAS ISOLATED FROM POOLED SERUM OF STRAIN  
 CC 13 INBRED GUINEA PIGS.  
 DR PTR: A02151; G2GP.  
 DR HSP: P01842; 7FAB.  
 DR InterPro: IPR003006; Ig\_MHC.  
 DR InterPro: IPR003597; Ig\_CL.  
 DR InterPro: IPR003600; Ig\_Like.  
 DR Pfam: PF00047; Ig; 2.  
 DR SMART: SM00410; IG\_Like; 1.  
 DR SMART: SM00407; IGCL; 2.  
 DR PROSITE: PS00290; IG\_MHC; 1.  
 KW Immunoglobulin domain; Immunoglobulin C region; Glycoprotein.  
 FT NON\_TER 1  
 FT DISULFID 16 16 INTERCHAIN (WITH A LIGHT CHAIN).  
 FT DISULFID 28 79  
 FT DISULFID 105 105 INTERCHAIN (WITH A HEAVY CHAIN).  
 FT DISULFID 107 107 INTERCHAIN (WITH A HEAVY CHAIN).  
 FT DISULFID 110 110 INTERCHAIN (WITH A HEAVY CHAIN).  
 FT DISULFID 142 202  
 FT CARBOHYD 178 178 N-LINKED (GLCNAC. . .).  
 FT DISULFID 248 308  
 SQ SEQUENCE 329 AA; 36074 MW; 5D231B7164D1FBA9 CRC64;  
 Query Match 67.0%; Score 894; DB 1; Length 329;  
 Best Local Similarity 68.8%; Pred. No. 4.2e-23;  
 Matches 165; Conservative 25; Mismatches 37; Indels 13; Gaps 2;  
 QY 1 MDKT-----HTCPCPAPPELLGGPSVFLFPPKPKDTLMISRTPEVTCVVDVSH 49  
 DB 90 VDKTVPRTZPBPCPCPCPPENLGGPSVFLFPPKPKDTLMISRTPEVTCVVDVSH 149  
 QY 50 EDPEKFNMYVDGVEVHNATKPREQYNSTYRVSVLTVLHODWLNGLKREYKCKVSNKAL 109  
 DB 150 DEPEQVTFWVDNKPVGNAETKPREQYNSTYRVSVLTVLHODWLNGLKREYKCKVSNKAL 209  
 QY 110 PAPIETISKAKQPREQVYTLPPSRDELTKNQVSLTCLVKGFYPSDIAVEWESNGQP- 168  
 DB 210 PAPIETISKAKQPREQVYTLPPSRDELTKNQVSLTCLVKGFYPSDIAVEWESNGQP- 269  
 QY 169 -ENNYKTTTPVLDSDGDFLYSKLVKRVDSKRMQGNVFCVSMHEALHNHYTKLSLSPG 227  
 DB 270 SEKEYNTPTIEDADGDFLYSKLVKRVDSKRMQGNVFCVSMHEALHNHYTKLSLSPG 329  
 RESULT 7  
 GC3\_MOUSE STANDARD; PRT; 329 AA.  
 ID GC3\_MOUSE  
 AC P22436;  
 DT 01-AUG-1991 (Rel. 19, Created)  
 DT 01-AUG-1991 (Rel. 19, Last sequence update)  
 DT 16-OCT-2001 (Rel. 40, Last annotation update)  
 DE Ig gamma-3 chain C region, secreted form.  
 OS Mus musculus (Mouse).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 OX NCBI\_TaxID=10090;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=85027161; PubMed=6092053;  
 RA Wels J.A., Word C.J., Rimm D., Der-Balan G.P., Martinez H.M.,  
 RA Tucker P.W., Blattner F.R.;  
 RT "Structural analysis of the murine IgG3 constant region gene.";  
 RL EMBO J. 3:2041-2046(1984).  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).





[2]  
 RN SEQUENCE OF 88-115 FROM N.A.  
 RC TISSUE-Fetal liver;  
 RX MEDLINE=83001943; PubMed=6811139;  
 RA Takahashi N., Oeda S., Ohta M., Nikaido T., Nakai S., Honjo T.;  
 RT "Structure of human immunoglobulin gamma genes: implications for  
 RL evolution of a gene family."; Cell 29:671-679(1982).  
 RN [3]  
 RN SEQUENCE OF 99-177 AND 310-326 FROM N.A.  
 RC TISSUE-Fetal liver;  
 RX MEDLINE=84235992; PubMed=6329676;  
 RA Krawinkel U., Rabbitts T.H.;  
 RT "Comparison of the hinge-coding segments in human immunoglobulin gamma  
 RL heavy chain genes and the linkage of the gamma 2 and gamma 4 subclass  
 RN genes."; EMBO J. 1:403-407(1982).  
 RN [4]  
 RN SEQUENCE OF 1-325 (MYELOMA PROTEIN TIL).  
 RX MEDLINE=81007873; PubMed=6774012;  
 RA Wang A.-C., Tung E., Fudenberg H.H.;  
 RT "The primary structure of a human IgG2 heavy chain: genetic,  
 RL evolutionary, and functional implications."; J. Immunol. 125:1048-1054(1980).  
 RN [5]  
 RN SEQUENCE OF 1-85 AND 132-325 (MYELOMA PROTEIN ZIE).  
 RX MEDLINE=80001357; PubMed=113060;  
 RA Connell G.E., Parr D.M., Hofmann T.;  
 RT "The amino acid sequences of the three heavy chain constant region  
 RL domains of a human IgG2 myeloma protein."; Can. J. Biochem. 57:758-767(1979).  
 RN [6]  
 RN SEQUENCE OF 238-275 (ZIE).  
 RX MEDLINE=80114419; PubMed=118920;  
 RA Hofmann T., Parr D.M.;  
 RT "A note of the amino acid sequence of residues 381-391 of human  
 RL immunoglobulins gamma chains."; Mol. Immunol. 16:923-925(1979).  
 RN [7]  
 RN REVISIONS TO 25; 59; 60 AND 264-268 (ZIE).  
 RX Hofmann T., Parr D.M.;  
 RL Submitted (MAR-1980) to the PIR data bank.  
 RN [8]  
 RN SEQUENCE OF 1-121 (DOT).  
 RX MEDLINE=95255298; PubMed=7737190;  
 RA Scoppini M., Bellotti V., Negri A., Merlini G., Garver F., Ferri G.;  
 RT "Characterization of the two unique human anti-flavin monoclonal  
 RL immunoglobulins."; Eur. J. Biochem. 228:886-893(1995).  
 RN [9]  
 RN DISULFIDE BONDS.  
 RX MEDLINE=72033500; PubMed=4940472;  
 RA Milstein C., Frangione B.;  
 RT "Disulphide bridges of the heavy chain of human immunoglobulin G2."; Biochem. J. 121:217-225(1971).  
 RN [10]  
 RN DISULFIDE BONDS.  
 RX MEDLINE=69064124; PubMed=5782707;  
 RA Frangione B., Milstein C., Fink J.R.L.;  
 RT "Structural studies of immunoglobulin G."; Nature 221:145-148(1969).  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL; J00230; AAB59393.1;  
 DR GSI; A02148; G2HU.  
 DR HSP; P01857; 1FC1.

DR Genew; HGNC:5526; IGHG2.  
 DR TIM; 147110;  
 DR InterPro: IPR003006; Ig\_MHC.  
 DR InterPro: IPR003597; Ig\_GL.  
 DR InterPro: IPR003600; Ig\_Like.  
 DR Pfam: PF00047; Ig; 3.  
 DR SMART: SM00410; Ig\_Like; 1.  
 DR SMART: SM00407; IG1; 2.  
 DR PROSITE: PS00290; Ig\_MHC; 2.  
 KW Immunoglobulin domain; Immunoglobulin C region.  
 FT NON\_TER 1 1  
 FT DOMAIN 1 98 CH1.  
 FT DOMAIN 99 110 HINGE.  
 FT DOMAIN 111 219 CH2.  
 FT DOMAIN 220 326 CH3.  
 FT DISULFID 14 14 INTERCHAIN (WITH A LIGHT CHAIN).  
 FT DISULFID 27 83  
 FT DISULFID 102 102 INTERCHAIN (WITH A HEAVY CHAIN).  
 FT DISULFID 103 103 INTERCHAIN (WITH A HEAVY CHAIN).  
 FT DISULFID 106 106 INTERCHAIN (WITH A HEAVY CHAIN).  
 FT DISULFID 109 109 INTERCHAIN (WITH A HEAVY CHAIN).  
 FT DISULFID 140 200  
 FT DISULFID 246 304  
 FT SITE 156 156  
 FT MOD\_RES 326 326  
 FT VARIANT 60 60  
 FT CONFLICT 109 109  
 FT SEQUENCE 326 AA; 35884 MW; 8310878C6878CF9C CRC64;  
 SQ  
 Query Match 85.9%; Score 1146; DB 1; Length 326;  
 Best Local Similarity 90.6%; Pred. No. 1e-31;  
 Matches 212; Conservative 9; Mismatches 6; Indels 7; Gaps 2;  
 QY 1 MDKT-----HTCPCPAPPELLGGPSVFLPPKPKDTLMISRTPEVTCVVVDVSHEDPEV 54  
 Db VDKTVERKCCVEGPPCPAPP-VAGPSVFLPPKPKDTLMISRTPEVTCVVVDVSHEDPEV 152  
 QY 55 KENNYVDGVEVHNAKTPREEQYNSTYRVSVLTVLDHQMNGKEYKCKVSNKALPAPIE 114  
 Db :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||: 212  
 QY 153 QFNWYVDGVEVHNAKTPREEQNFSTRVSVLTVLDHQMNGKEYKCKVSNKALPAPIE 212  
 Db :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||: 174  
 QY 115 KTISKARGQPREQVYTLPPSRDELTKNQVSLTCLVKGFYPSDIAVEMESNGQPENNYKT 272  
 Db :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||: 272  
 QY 213 KTISKARGQPREQVYTLPPSRDELTKNQVSLTCLVKGFYPSDIAVEMESNGQPENNYKT 272  
 Db :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||: 228  
 QY 175 TTPVLDSDGSEFFLYSKLTVDKSRWQQGNVSCFVSVMHEALHNHYTQKSLSLSPGK 228  
 Db :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||: 326  
 QY 273 TTPVLDSDGSEFFLYSKLTVDKSRWQQGNVSCFVSVMHEALHNHYTQKSLSLSPGK 326  
 Db :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:  
 RESULT 3  
 GC4\_HUMAN  
 ID GC4\_HUMAN STANDARD; PRT; 327 AA.  
 AC P01861;  
 DT 21-JUL-1986 (Rel. 01, Created)  
 DT 21-JUL-1986 (Rel. 01, Last sequence update)  
 DT 16-OCT-2001 (Rel. 40, Last annotation update)  
 DE Ig gamma-4 chain C region.  
 GN IGHG4.  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 OX NCBI\_Taxid=9606;  
 RN [1]  
 RN SEQUENCE FROM N.A.  
 RX MEDLINE=83157104; PubMed=6299662;  
 RA Ellison J.W., Buxbaum J.N., Hood L.E.;  
 RT "Nucleotide sequence of a human immunoglobulin C gamma 4 gene."; DNA 1:11-18(1981).  
 RL [2]  
 RN SEQUENCE OF 1-30 AND 81-326.  
 RX MEDLINE=70207560; PubMed=4192699;

RT Intrachain disulfide bonds.;  
 RL Biochemistry 9:3188-3196(1970).  
 [7]  
 RX MEDLINE=77070267; PubMed=1002129;  
 RA DISULFIDE BONDS.  
 RA Dreker L., Schwarz J., Reichel W., Hilschmann N.;  
 "Rule of antibody structure. The primary structure of a monoclonal  
 RT IgG1 immunoglobulin (myeloma protein Nie), I: Purification and  
 RT characterization of the protein, the L- and H-chains, the  
 RT cyanogen bromide cleavage products, and the disulfide bridges.";  
 RL Hoppe-Seyler's Z. Physiol. Chem. 357:1515-1540(1976).  
 [8]  
 RP X-RAY CRYSTALLOGRAPHY (2.9 ANGSTROMS).  
 RA MEDLINE=81208100; PubMed=7236608;  
 RA Deisenhofer J.;  
 RT "Crystallographic refinement and atomic models of a human Fc fragment  
 RT and its complex with fragment B of protein A from Staphylococcus  
 RL aureus at 2.9- and 2.8-A resolution.";  
 CC Biochemistry 20:2361-2370(1981).  
 CC -!- MISCELLANEOUS: NIE HAS THE GLM(17) ALLOTYPIC MARKER, 97-K, & THE  
 CC GLM(1) MARKERS, 239-D & 241-L. KOL & EU SEQUENCES HAVE THE GLM(3)  
 CC MARKER & THE GLM (NON-1) MARKERS.  
 CC -!- MISCELLANEOUS: NIE ALSO DIFFERS IN THE AMIDATION STATES OF  
 CC 35,116,198,289 & 272.  
 CC -!- MISCELLANEOUS: EU ALSO DIFFERS IN THE AMIDATION STATES OF RESIDUES  
 CC 155, 166, 177, 195, 198, 269, AND 272 AND IN THE ORDER OF RESIDUES  
 CC 268-272.  
 CC -!- MISCELLANEOUS: KOL ALSO DIFFERS IN THE AMIDATION STATES OF  
 CC RESIDUES 198,267&272.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 CC EMBL; J00228; AAC82527.1; ALT\_INIT.  
 DR PIR; A02146; GHU.  
 DR PDB; 1FC1; 15-JUL-92.  
 DR PDB; 1FC2; 15-JUL-92.  
 DR Genew; HGNC:5525; IGHG1.  
 DR MIM; 147100; -  
 DR InterPro; IPR003006; Ig\_MHC.  
 DR InterPro; IPR003597; Ig\_c1.  
 DR InterPro; IPR003600; Ig\_like.  
 DR Pfam; PF00047; Ig; 3.  
 DR SMART; SM00410; Ig\_like; 1.  
 DR SMART; SM00407; IGHG1; 2.  
 DR PROSITE; PS00290; Ig\_MHC; 2.  
 KW Immunoglobulin domain; Immunoglobulin C region; Glycoprotein;  
 KW 3D-structure.  
 FT NON\_TER 1 1  
 FT DOMAIN 1 98  
 FT CH1.  
 FT HINGE.  
 FT DOMAIN 99 110  
 FT CH2.  
 FT DOMAIN 111 223  
 FT CH3.  
 FT DOMAIN 224 330  
 FT DISULFID 27 83  
 FT DISULFID 103 103  
 FT DISULFID 109 109  
 FT DISULFID 112 112  
 FT DISULFID 144 204  
 FT DISULFID 250 308  
 FT DISULFID 180 180  
 FT CARBOHYD 330 330  
 FT MOD\_RES 97 97  
 FT VARIANT 239 239  
 FT VARIANT 241 241  
 FT VARIANT 241 241  
 FT STRAND 123 126

FT HELIX 130 134  
 FT TURN 136 137  
 FT STRAND 141 148  
 FT STRAND 158 162  
 FT TURN 163 164  
 FT STRAND 165 166  
 FT STRAND 175 178  
 FT STRAND 183 190  
 FT STRAND 193 197  
 FT HELIX 198 199  
 FT TURN 202 206  
 FT STRAND 215 219  
 FT STRAND 227 227  
 FT STRAND 230 234  
 FT STRAND 238 240  
 FT TURN 241 242  
 FT STRAND 245 256  
 FT STRAND 260 266  
 FT TURN 267 268  
 FT STRAND 274 276  
 FT STRAND 280 281  
 FT TURN 283 284  
 FT STRAND 287 296  
 FT HELIX 297 301  
 FT TURN 302 303  
 FT STRAND 306 312  
 FT TURN 313 314  
 FT TURN 316 317  
 FT STRAND 320 324  
 SQ SEQUENCE 330 AA; 36106 MW; 3770EE106C2FA33D CRC64;  
 Query Match 92.4%; Score 1233; DB 1; Length 330;  
 Best Local Similarity 100.0%; Pred. No. 1.1e-34;  
 Matches 227; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 2 DKHTCPPCPAPELGGPSVFLPPPKPDTLMISRTPEVTCVVVDVSHEDPEVKFNWYVD 61  
 Db 104 DKHTCPPCPAPELGGPSVFLPPPKPDTLMISRTPEVTCVVVDVSHEDPEVKFNWYVD 163  
 QY 62 GVEVHNATKPREQYNSTYRVSVLTVLDHQLNCKEYKCKVSNKALPAPIETKTISKAK 121  
 Db 164 GVEVHNATKPREQYNSTYRVSVLTVLDHQLNCKEYKCKVSNKALPAPIETKTISKAK 223  
 QY 122 GQPREPQVYTLPPSRDELTKNQVSLTCLVKGFYPSDIAVEWESNGQPENNYKTTPPVLD 181  
 Db 224 GQPREPQVYTLPPSRDELTKNQVSLTCLVKGFYPSDIAVEWESNGQPENNYKTTPPVLD 283  
 QY 182 DGSFFLYSKLTVDKSRWQQGNVSCVMHEALHNHYTQKSLSLSPGK 228  
 Db 284 DGSFFLYSKLTVDKSRWQQGNVSCVMHEALHNHYTQKSLSLSPGK 330  
 RESULT 2  
 GC2\_HUMAN  
 ID GC2\_HUMAN STANDARD; PRT; 326 AA.  
 AC P01859;  
 DT 21-JUL-1986 (Rel. 01, Created)  
 DT 21-JUL-1986 (Rel. 01, Last sequence update)  
 DT 16-OCT-2001 (Rel. 40, Last annotation update)  
 DE Ig gamma-2 chain C region.  
 OS IGHG2.  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Euthera; Primates; Catarrhini; Hominidae; Homo.  
 OX NCBI\_Taxid=9606;  
 RN [1]  
 RP SEQUENCE OF 2-326 FROM N.A.  
 RX MEDLINE=82197621; PubMed=6804948;  
 RA Ellison J.W., Hood L.E.;  
 RT "Linkage and sequence homology of two human immunoglobulin gamma  
 RT heavy chain constant region genes";  
 RL Proc. Natl. Acad. Sci. U.S.A. 79:1984-1988(1982).

GenCore version 5.1.4\_p5\_4578  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: April 21, 2003, 10:42:21 ; Search time 9.40588 seconds  
(without alignments)  
1084.766 Million cell updates/sec

Title: 2LINK7

Perfect score: 1334

Sequence: 1 MDKTHTCPPCAPPELLGGPS.....GKDLKAFYDKVAEKLEAF 246

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.0

Searched: 112892 seqs, 41476328 residues

Total number of hits satisfying chosen parameters: 112892

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt\_40.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Query %	Score	Match	Length	ID	Description
1	1233	92.4	330	1	GCL_HUMAN	P01857 homo sapien
2	1146	85.9	326	1	GC2_HUMAN	P01859 homo sapien
3	1142	85.6	327	1	GC4_HUMAN	P01861 homo sapien
4	1138	85.3	290	1	GC3_HUMAN	P01860 homo sapien
5	921	69.0	323	1	GC_RABIT	P01870 oryctolagus
6	894	67.0	329	1	GC2_CAVPO	P01862 caria porce
7	848	63.6	329	1	GC3_MOUSE	P22436 mus musculu
8	840	63.0	333	1	GC8_RAT	P20761 rattus norv
9	838	62.8	398	1	GC3M_MOUSE	P03987 mus musculu
10	825	61.8	326	1	GC1_RAT	P20759 rattus norv
11	823	61.7	324	1	GC1_MOUSE	P01868 mus musculu
12	823	61.7	393	1	GC1M_MOUSE	P01869 mus musculu
13	819	61.4	399	1	GCAM_MOUSE	P01865 mus musculu
14	818	61.3	329	1	GCC_RAT	P20762 rattus norv
15	813	60.9	330	1	GCAA_MOUSE	P01863 mus musculu
16	809	60.6	335	1	GCAH_MOUSE	P01864 mus musculu
17	793	59.4	322	1	GC8_RAT	P20760 rattus norv
18	791	59.3	405	1	GC8M_MOUSE	P01867 mus musculu
19	785	58.8	336	1	GC8_MOUSE	P01866 mus musculu
20	387	28.0	421	1	EPC_MOUSE	P06336 mus musculu
21	382	28.6	454	1	MUC_HUMAN	P01871 homo sapien
22	372	27.9	429	1	EPC_RAT	P01855 rattus norv
23	370	27.7	391	1	MUCB_HUMAN	P04220 homo sapien
24	370	27.7	455	1	MUC_MOUSE	P01872 mus musculu
25	369	27.7	428	1	EPC_HUMAN	P01854 homo sapien
26	367	27.5	458	1	MUC_RABIT	P03988 oryctolagus
27	364	27.3	476	1	MUCM_MOUSE	P01873 mus musculu
28	361	27.1	479	1	MUCM_RABIT	P04221 oryctolagus
29	359	26.7	454	1	MUC_MESAU	P06337 mesocricetu
30	346	26.2	450	1	MUC_CANFA	P01874 canis fami
31	347	26.0	457	1	MUC_SUNMU	P20768 suncus muri
32	326	24.4	446	1	MUC_CHICK	P01875 gallus gall
33	321	24.1	438	1	HVC2_HETFR	P23085 heterodontu

34	317	23.8	6486	1	TYCC_BACBR	O30409 b tyrocidin
35	313	23.5	438	1	HVCS_HETFR	P23087 heterodontu
36	312	23.4	5037	1	RYR1_RABIT	P11716 oryctolagus
37	312	23.4	5038	1	RYR1_HUMAN	P21817 homo sapien
38	311	23.3	461	1	HVCM_HETFR	P23088 heterodontu
39	310	23.2	4969	1	RYR2_RABIT	P30957 oryctolagus
40	308	23.1	299	1	ALC_RABIT	P01879 oryctolagus
41	307	23.0	4725	1	DYHC_DICDI	P34036 dictyostelli
42	306	22.9	4367	1	DYHC_NEUCR	P45443 neurospora
43	306	22.9	4367	1	RYR2_HUMAN	Q92736 homo sapien
44	305	22.9	370	1	HVC1_HETFR	P23084 heterodontu
45	305	22.9	4393	1	PGBM_HUMAN	P98160 homo sapien

## ALIGNMENTS

RESULT 1  
GCL\_HUMAN STANDARD; PRT; 330 AA.  
AC P01857:  
DT 21-JUL-1986 (Rel. 01, Created)  
DT 21-JUL-1986 (Rel. 01, Last sequence update)  
DT 16-OCT-2001 (Rel. 40, Last annotation update)  
DE Ig gamma-1 chain C region.  
GN IGHG1.  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
OX NCBI\_TaxID=9606;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=82274238; PubMed=6287432;  
RA Ellison J.W., Berson B.J., Hood L.E.;  
RT "The nucleotide sequence of a human Immunoglobulin C gamma gene.";  
RL Nucleic Acids Res. 10:4071-4079(1982).  
RN [2]  
RP SEQUENCE OF 1-135 (MYELOMA PROTEIN EU).  
RX MEDLINE=71064024; PubMed=5489771;  
RA Cunningham B.A., Rutishauser U., Gall W.E., Gottlieb P.D.,  
RA Waxdal M.J., Edelman G.M.;  
RT "The covalent structure of a human gamma G-immunoglobulin. VII. Amino acid sequence of heavy-chain cyanogen bromide fragments H1-H4.";  
RN [3]  
RP SEQUENCE OF 136-329 (EU).  
RX MEDLINE=71064025; PubMed=5530842;  
RA Rutishauser U., Cunningham B.A., Bennett C., Konigsberg W.H.,  
RA Edelman G.M.;  
RT "The covalent structure of a human gamma G-immunoglobulin. 8. Amino acid sequence of heavy-chain cyanogen bromide fragments H5-H7.";  
RN [4]  
RP SEQUENCE (MYELOMA PROTEIN NIE).  
RX MEDLINE=77070269; PubMed=826475;  
RA Ponstingl H., Hilschmann N.;  
RT "The rule of antibody structure. The primary structure of a monoclonal IgG1 immunoglobulin (myeloma protein Nie). III. The chymotryptic peptides of the H-chain, alignment of the tryptic peptides and discussion of the complete structure.";  
RL Hoppe-Seyler's Z. Physiol. Chem. 357:1571-1604(1976).  
RN [5]  
RP SEQUENCE (MYELOMA PROTEIN KOL), AND DISULFIDE BONDS.  
RX MEDLINE=83289131; PubMed=6884994;  
RA Schmidt W.E., Jung H.-D., Palm W., Hilschmann N.;  
RT "Three-dimensional structure determination of antibodies. Primary structure of crystallized monoclonal immunoglobulin IgG1 KOL, I.";  
RL Hoppe-Seyler's Z. Physiol. Chem. 364:713-747(1983).  
RN [6]  
RP DISULFIDE BONDS.  
RX MEDLINE=71064027; PubMed=4923144;  
RA Gall W.E., Edelman G.M.;  
RT "The covalent structure of a human gamma G-immunoglobulin. X.







Db 151 FTWYINNEQVARTPLREQQFNSTIRVSTLTPIHQDWLRGKFKCKVHNKALPAPIEK 210  
QY 116 TISKAKGQPREQVYTLPLPSRDELTKNOVSLTCLVKGYFSPDSIAEVESNGQPENNYKTT 175  
Db 211 TISKARGQPLEPKVYTMGPPEELSSRSVSLFCMNGFFPSDISVEWENKRAEDNYKTT 270  
QY 176 PVLVSDSGSFFLYSKLTVDKSRWQGNVFCSSVMHEALHNHYTKQSLSLSPGK 228  
Db 271 PAVLSDSGSYFLYNKLSVPTSEWQGDVFTCSVMHEALHNHYTKQSLSPGK 323

## RESULT 11

I47160  
Ig gamma 2b chain constant region - pig (fragment)  
C:Species: Sus scrofa domestica (domestic pig)  
C:Date: 21-Feb-1997 #sequence\_revision 21-Feb-1997 #text\_change 21-Jan-2000  
C:Accession: I47160  
R:Kaskovics, I.; Sun, J.; Butler, J.E.  
J. Immunol. 153, 3565-3573, 1994  
A:Title: Five putative subclasses of swine IgG identified from the cDNA sequences of a  
A:Reference number: I47158; MUID:95015845; PMID:7930579  
A:Accession: I47160  
A:Status: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: mRNA  
A:Residues: 1-328 <KAC>  
A:Cross-references: EMBL:U03780; NID:9433125; PIDN:AAA52218.1; PID:9433126  
C:Genetics:  
A:Gene: IgG2b  
C:Superfamily: immunoglobulin C region; immunoglobulin homology  
F:133-202/Domain: immunoglobulin homology <IMM>

Query Match 68.7%; Score 917; DB 2; Length 328;  
Best Local Similarity 65.4%; Pred. No. 1.4e-23;  
Matches 166; Conservative 32; Mismatches 24; Indels 32; Gaps 4;

QY 3 KTHTC-----PPCP-----APELLGSPSVFLPPPKDITMISR 36  
Db 79 KSYTCNVNHPATTTKVDRKVGTKTPPCPCIPACESP-----GPSVFIFPPKPKDITMISR 134  
QY 37 TPEVTCVVVDVSHEDPEVKFNMYVDGVEVHNKATKPREQYNSTYRVVSVLTVLHQDWLNL 96  
Db 135 TPQVTCVVVDVSDGVNQFNSVVDGVEVHTAQTTRKPEQFNSTYRVVSVLPIQHODWLN 194  
QY 97 GREYKCKVSNKALPAPIEKTISKAKGQPREPOVYTLPPSRDELTKNOVSLTCLVKGYFYS 156  
Db 195 GREFKCKVNNKDLPAITRIISKAGQTRPQVYTLPPHAELSRKSVITCLVIGFYPP 254  
QY 157 DIAVESNGQ--PENNYKTTTPVLSDSGSFFLYSKLTVDKSRWQGNVFCSSVMHEALH 214  
Db 255 DIDVEMQRNGQPEPEGNRYRTTPQDDVDGTGYFLYSKFSVDKASWQGGGIFQCAVMHEALH 314  
QY 215 NHYTKSLSLSPGK 228  
Db 315 NHYTKSISKTGPK 328

## RESULT 12

I47159  
Ig gamma 2a chain constant region - pig (fragment)  
C:Species: Sus scrofa domestica (domestic pig)  
C:Date: 21-Feb-1997 #sequence\_revision 21-Feb-1997 #text\_change 21-Jan-2000  
C:Accession: I47159  
R:Kaskovics, I.; Sun, J.; Butler, J.E.  
J. Immunol. 153, 3565-3573, 1994  
A:Title: Five putative subclasses of swine IgG identified from the cDNA sequences of a  
A:Reference number: I47158; MUID:95015845; PMID:7930579  
A:Accession: I47159  
A:Status: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: mRNA  
A:Residues: 1-328 <KAC>  
A:Cross-references: EMBL:U03779; NID:9433123; PIDN:AAA52217.1; PID:9433124  
C:Genetics:  
A:Gene: IgG2a

C:Superfamily: immunoglobulin C region; immunoglobulin homology  
F:133-202/Domain: immunoglobulin homology <IMM>

Query Match 68.7%; Score 917; DB 2; Length 328;  
Best Local Similarity 65.4%; Pred. No. 1.4e-23;  
Matches 166; Conservative 32; Mismatches 24; Indels 32; Gaps 4;

QY 3 KTHTC-----PPCP-----APELLGSPSVFLPPPKDITMISR 36  
Db 79 KSYTCNVNHPATTTKVDRKVGTKTPPCPCIPACESP-----GPSVFIFPPKPKDITMISR 134  
QY 37 TPEVTCVVVDVSHEDPEVKFNMYVDGVEVHNKATKPREQYNSTYRVVSVLTVLHQDWLNL 96  
Db 135 TPQVTCVVVDVSDGVNQFNSVVDGVEVHTAQTTRKPEQFNSTYRVVSVLPIQHODWLN 194  
QY 97 GREYKCKVSNKALPAPIEKTISKAKGQPREPOVYTLPPSRDELTKNOVSLTCLVKGYFYS 156  
Db 195 GREFKCKVNNKDLPAITRIISKAGQTRPQVYTLPPHAELSRKSVITCLVIGFYPP 254  
QY 157 DIAVESNGQ--PENNYKTTTPVLSDSGSFFLYSKLTVDKSRWQGNVFCSSVMHEALH 214  
Db 255 DIDVEMQRNGQPEPEGNRYRTTPQDDVDGTGYFLYSKFSVDKASWQGGGIFQCAVMHEALH 314  
QY 215 NHYTKSLSLSPGK 228  
Db 315 NHYTKSISKTGPK 328

## RESULT 13

I47162  
Ig gamma 4 chain constant region - pig (fragment)  
C:Species: Sus scrofa domestica (domestic pig)  
C:Date: 21-Feb-1997 #sequence\_revision 21-Feb-1997 #text\_change 21-Jan-2000  
C:Accession: I47162  
R:Kaskovics, I.; Sun, J.; Butler, J.E.  
J. Immunol. 153, 3565-3573, 1994  
A:Title: Five putative subclasses of swine IgG identified from the cDNA sequences of  
A:Reference number: I47158; MUID:95015845; PMID:7930579  
A:Accession: I47162  
A:Status: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: mRNA  
A:Residues: 1-277 <KAC>  
A:Cross-references: EMBL:U03782; NID:9433129; PIDN:AAA52220.1; PID:9433130  
C:Genetics:  
A:Gene: IgG4  
C:Superfamily: immunoglobulin C region; immunoglobulin homology  
F:82-151/Domain: immunoglobulin homology <IMM>

Query Match 68.3%; Score 911; DB 2; Length 277;  
Best Local Similarity 66.4%; Pred. No. 1.3e-23;  
Matches 166; Conservative 31; Mismatches 29; Indels 24; Gaps 4;

QY 3 KTHTC-----PPCP-----APELLG-----GPSVFIFPPKPKDITMISRTPEV 40  
Db 28 KSYTCNVNHPATTTKVDRKVGTKTPPCPCIPACEGGPSAFIFPPKPKDITMISRTPKV 87  
QY 41 TCVVVDVSHEDPEVKFNMYVDGVEVHNKATKPREQYNSTYRVVSVLTVLHQDWLNGKEY 100  
Db 88 TCVVVDVSDGVNQFNSVVDGVEVHTAQTTRKPEQFNSTYRVVSVLPIQHODWLNKEF 147  
QY 101 KCKVSNKALPAPIEKTISKAKGQPREPOVYTLPPSRDELTKNOVSLTCLVKGYFSPDIAY 160  
Db 148 KCKVNNKDLPAITRIISKAGQTRPQVYTLPPPTTEELSRKSVITCLVIGFYPPDIDV 207  
QY 161 EWESNGQ--PENNYKTTTPVLSDSGSFFLYSKLTVDKSRWQGNVFCSSVMHEALHNHYT 218  
Db 208 EWQRNGQPEPEGNRYRTTPQDDVDGTGYFLYSKLAVDKASWQRGDTFQCAVMHEALHNHYT 267  
QY 219 QKSLSLSPGK 228  
Db 268 QKSIFKTPGK 277

## RESULT 9

G3HUW1  
 Ig gamma-3 heavy chain disease proteins - human  
 C:Species: Homo sapiens (man)  
 C:Date: 31-Dec-1979 #sequence\_revision 23-Oct-1981 #text\_change 16-Jul-1999  
 C:Accession: A90442; A92219; A90198; A93915; A02149  
 R:Frangione, B.; Rosenwasser, E.; Prelli, F.; Franklin, E.C.  
 Biochemistry 19, 4304-4308, 1980  
 A:Title: Primary structure of human gamma3 immunoglobulin deletion mutant: gamma3 heavy-  
 A:Reference number: A90442; MUID:81021548; PMID:6774747  
 A:Contents: heavy chain disease protein Wis  
 A:Accession: A90442  
 A:Molecule type: protein  
 A:Residues: 1-289 <FRA>  
 A:Note: the molecule is a dimer linked by 12 disulfide bonds: it has an extra interchain  
 A:Note: this protein lacks most of the V region and all of the CH1 region. Residue 12 co  
 A:Note: the sequence of residues 42-76 was taken from the reference that follows  
 R:Michaelson, T.E.; Frangione, B.; Franklin, E.C.  
 J. Biol. Chem. 252, 883-889, 1977  
 A:Title: Primary structure of the 'hinge' region of human IgG3. Probable quadruplication  
 A:Reference number: A92219; MUID:77118561; PMID:402363  
 A:Contents: normal gamma-3 chains, sequence corresponding to residues 12-97 of protein W  
 A:Accession: A92219  
 A:Molecule type: protein  
 A:Residues: 12-97 <MIC>  
 A:Note: the hinge region in gamma-3 chains is about four times as long as in other gamma  
 A:Note: cysteines at positions 24, 27, 33, 39, 42, 48, 54, 57, 63, 69, and 72 form inter  
 R:Wolfenstein-Todel, C.; Frangione, B.; Prelli, F.; Franklin, E.C.  
 Biochem. Biophys. Res. Commun. 71, 907-914, 1976  
 A:Title: The amino acid sequence of "heavy chain disease" protein ZUC. Structure of the  
 A:Reference number: A90198; MUID:77021516; PMID:823945  
 A:Contents: heavy chain disease protein Zuc, partial sequence corresponding to residues  
 A:Accession: A90198  
 A:Molecule type: protein  
 A:Residues: 59-125, 'EB', 128-226, 228-289 <WOL>  
 A:Note: this protein lacks most of the V region, all of the CH1 region, and part of the  
 R:Alexander, A.; Steinmetz, M.; Barritault, D.; Frangione, B.; Franklin, E.C.; Hood, L.;  
 Proc. Natl. Acad. Sci. U.S.A. 79, 3260-3264, 1982  
 A:Title: gamma heavy chain disease in man: CDNA sequence supports partial gene deletion  
 A:Reference number: A93915; MUID:82247835; PMID:6808505  
 A:Contents: heavy chain disease protein Omn  
 A:Accession: A93915  
 A:Molecule type: mRNA  
 A:Residues: 12-70; 72-114; 116-125, 'E', 127-133, 'L', 135-136, 'E', 138, 'Y', 140-154, 'D', 156-157  
 A:Note: a carboxyl-terminal Lys is removed posttranslationally  
 A:Note: this sequence may represent an allelic form or another gamma chain subclass  
 C:Comment: The heavy chain disease protein Wis is shown.  
 C:Genetics:  
 A:Gene: GDB:IGHG3  
 A:Cross-references: GDB:119339; OMIM:147120  
 A:Map position: 14q32.33-14q32.33  
 C:Superfamily: immunoglobulin C region; immunoglobulin homology  
 C:Keywords: duplication; glycoprotein; immunoglobulin; pyroglyutamic acid  
 F:203-270/Domain: immunoglobulin homology <IMM>  
 F:1/Modified site: pyrrolidone carboxylic acid (Gln) #status experimental  
 F:6,140/Binding site: carbohydrate (Asn) (covalent) #status experimental  
 Query Match 84.9%; Score 1133; DB 1; Length 289;  
 Best Local Similarity 76.4%; Pred. No. 3.8e-31;  
 Matches 207; Conservative 13; Mismatches 6; Indels 45; Gaps 1;

QY 2 DKTHT-----  
 Db 19 DTHHTCRCEPKSCDTPPPCPCPEPKSCDTPPPCPCPEPKSCDTPPPCPCPEPELL 78  
 QY 17 GGSVFLFPKPKDLMISRTPEVTCVVDVSHDEPEVFNWYDGVVEHNAKTKPREQ 76  
 Db 79 GGSVFLFPKPKDLMISRTPEVTCVVDVSHDEPEVFNWYDGVVEHNAKTKPREQ 138  
 QY 77 YNSTYRVSVLTVLHQNLDNGKEYCKVSNKALPAPIEKTISKAKGQPREPQVYTLPPSR 136  
 Db 139 FNSTFRVSVLTVLHQNLDNGKEYCKVSNKALPAPIEKTISKAKGQPREPQVYTLPPSR 198

QY 137 DELTKNOVSLTCLVKGFYPSDIAVWESNGQPENNYKTTTPPVLDSDGSGFFLYSKLTVDKS 196  
 Db 199 EEMTKNOVSLTCLVKGFYPSDIAVWESNGQPENNYKTTTPPVLDSDGSGFFLYSKLTVDKS 258  
 QY 197 RWQOGNVFSCSVMHAEALHNHYTKSLSLSPG 227  
 Db 259 RWQOGNIFSCSVMHAEALHNRYTKSLSLSPG 289

## RESULT 10

## GHRB

Ig gamma chain C region - rabbit  
 C:Species: Oryctolagus cuniculus (domestic rabbit)  
 C:Date: 24-Apr-1984 #sequence\_revision 15-Nov-1984 #text\_change 16-Jul-1999  
 C:Accession: A91749; A90290; A93928; A90245; A94416; A02161  
 R:Bernstein, K.E.; Alexander, C.B.; Mage, R.G.  
 Immunogenetics 18, 387-397, 1983  
 A:Title: Nucleotide sequence of a rabbit IgG heavy chain from the recombinant F-1a1  
 A:Reference number: A91749; MUID:84030930; PMID:6313520  
 A:Accession: A91749  
 A:Molecule type: mRNA  
 A:Residues: 1-323 <BER>  
 A:Note: this sequence has the d12 allotypic marker, 104-Thr, and the e14 marker, 185-  
 R:Pratt, D.M.; Mole, L.E.  
 Biochem. J. 151, 337-349, 1975  
 A:Title: Sequence studies on the constant region of the Fd sections of rabbit immuno-  
 A:Reference number: A90290; MUID:76135469; PMID:1243651  
 A:Accession: A90290  
 A:Molecule type: protein  
 A:Residues: 1-47, 'E', 49-71, 'pv', 72-128 <PRA>  
 R:Martens, C.L.; Moore, K.W.; Steinmetz, M.; Hood, L.; Knight, K.L.  
 Proc. Natl. Acad. Sci. U.S.A. 79, 6018-6022, 1982  
 A:Title: Heavy chain genes of rabbit IgG; isolation of a cDNA encoding gamma heavy ci  
 A:Reference number: A93928; MUID:83299917; PMID:6193512  
 A:Accession: A93928  
 A:Molecule type: mRNA  
 A:Residues: 88-103, 'M', 105-143, 'E', 145-184, 'A', 186, 'E', 188-266 <MAR>  
 A:Cross-references: GB:M16426; NID:gl65111; PIDN:AAA31289.1; PID:gl65112  
 A:Note: this sequence has the d11 allotypic marker, 104-Met, and the e15 allotypic m  
 R:Fruchter, R.G.; Jackson, S.A.; Moie, L.E.; Porter, R.R.  
 Biochem. J. 116, 249-259, 1970  
 A:Title: Sequence studies of the Fd section of the heavy chain of rabbit immunoglobu-  
 A:Reference number: A90245; MUID:70110015; PMID:5461106  
 A:Accession: A90245  
 A:Molecule type: protein  
 A:Residues: 132-143, 'E', 145-161 <FRU>  
 R:Hill, R.L.; Lebowitz, H.E.; Fellows Jr., R.E.; Delaney, R.  
 in Gamma Globulins, Nobel Symp. 3, Killander, J., ed., pp.109-127, Almqvist and Wikse  
 A:Reference number: A94416  
 A:Accession: A94416  
 A:Molecule type: protein  
 A:Residues: 129-131; 155-172, 'D', 174-184, 'A', 186, 'E', 188-200, 'D', 202-217, 'E', 219-232, '  
 A:Note: this has the e15 allotypic marker, 185-Ala  
 C:Complex: An immunoglobulin heterotetramer subunit consists of two identical light (C  
 hain disulfide bonds. In some cases, such as IgA and IgM, the subunits associate int  
 C:Superfamily: immunoglobulin C region; immunoglobulin homology  
 C:Keywords: duplication; glycoprotein; heterotetramer; immunoglobulin  
 F:20-82/Domain: immunoglobulin homology <IM1>  
 F:130-199/Domain: immunoglobulin homology <IM2>  
 F:236-303/Domain: immunoglobulin homology <IM3>  
 F:173/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 69.0%; Score 921; DB 1; Length 323;  
 Best Local Similarity 71.7%; Pred. No. 1e-23;  
 Matches 167; Conservative 29; Mismatches 32; Indels 5; Gaps 2;  
 QY 1 MDKT---HTC--PPCPAPELLGGPSVFLPPEPKDLMISRTPEVTCVVDVSHDEPEVK 55  
 Db 91 VDKTVAPTSCSRKTPPELLGGPSVFLPPEPKDLMISRTPEVTCVVDVSHDEPEVK 150  
 QY 56 FNNYDGVVEHNAKTKPREEQYNSTYRVSVLTVLHQNLDNGKEYCKVSNKALPAPIEK 115



QY 182 DGSFFLYSKLTVDKSRWQGNVFCSCVMHEALHNHYTKQSLSPGK 228  
 |||  
 Db 328 DGSFFLYSKLTVDKSRWQGNVFCSCVMHEALHNHYTKQSLSPGK 374

## RESULT 4

PT0207  
 Ig gamma chain C region - chimpanzee  
 C:Species: Pan troglodytes (chimpanzee)  
 C:Date: 23-Nov-1991 #sequence\_revision 23-Nov-1991 #text\_change 16-Jul-1999  
 C:Accession: PT0207  
 R: Ehrlich, P.H.; Moustafa, Z.A.; Oestberg, L.  
 Mol. Immunol. 28, 319-322, 1991  
 A:Title: Nucleotide sequence of chimpanzee Fc and hinge regions.  
 A:Reference number: PT0207; MUID:91287716; PMID:2062315  
 A:Accession: PT0207  
 A:Molecule type: mRNA  
 A:Residues: 1-234 <EHR>  
 C:Superfamily: immunoglobulin C region; immunoglobulin homology  
 C:Keywords: immunoglobulin  
 F:48-117/Domain: immunoglobulin homology <IMM>

Query Match 88.5%; Score 1180; DB 2; Length 234;  
 Best Local Similarity 98.6%; Pred. No. 4.6e-33;  
 Matches 217; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 2 DKTHTCPPCAPPELLGSPVFLFPKPKDTLMISRTPEVTCVVVDVSHEDPEVKFNWYVD 61  
 |||  
 Db 15 DTHTTCPPCAPPELLGSPVFLFPKPKDTLMISRTPEVTCVVVDVSHEDPEVKFNWYVD 74  
 |||  
 QY 62 GVEVHNAKTPREQNSTYRVSVTLVHODWLNQKEYCKVSNKALPAPIETISKAK 121  
 |||  
 Db 75 GVEVHNAKTPREQNSTYRVSVTLVHODWLNQKEYCKVSNKALPAPIETISKAK 134  
 |||  
 QY 122 GQPREQVYTLPPSRDELTKNQVSLTCLVKGFYPSDIAVEWESNGQPENNYKTPPVLD 181  
 |||  
 Db 135 GQPREQVYTLPPSRDELTKNQVSLTCLVKGFYPSDIAVEWESNGQPENNYKTPPVLD 194  
 |||  
 QY 182 DGSFFLYSKLTVDKSRWQGNVFCSCVMHEALHNHYTKQSLSPGK 221  
 |||  
 Db 195 DGSFFLYSKLTVDKSRWQGNVFCSCVMHEALHNHYTKQSLSPGK 234  
 |||

## RESULT 5

A23511  
 Ig gamma-3 chain C region (allotype G3m(b)) - human  
 C:Species: Homo sapiens (man)  
 C:Date: 28-Dec-1987 #sequence\_revision 28-Dec-1987 #text\_change 23-Jul-1999  
 C:Accession: A23511  
 R: Huck, S.; Fort, P.; Crawford, D.H.; Lefranc, M.P.; Lefranc, G.  
 Nucleic Acids Res. 14, 1779-1789, 1986  
 A:Title: Sequence of a human immunoglobulin gamma 3 heavy chain constant region gene: cDNA  
 A:Reference number: A23511; MUID:86148507; PMID:3081877  
 A:Accession: A23511  
 A:Molecule type: DNA  
 A:Residues: 1-377 <HUC>  
 A:Cross-references: GB:X03604; GB:M12958; NID:g33070; PIDN:CAA27268.1; PID:g577056  
 C:Genetics:  
 A:Gene: GDB:IGHG3  
 A:Cross-references: GDB:119339; OMIM:147120  
 A:Map position: 14q32.33-14q32.33  
 A:Introns: 98/3; 115/3; 130/3; 145/3; 160/3; 270/3  
 C:Superfamily: immunoglobulin C region; immunoglobulin homology  
 C:Keywords: immunoglobulin  
 F:20-85/Domain: immunoglobulin homology <IMM>

Query Match 86.8%; Score 1158; DB 2; Length 377;  
 Best Local Similarity 78.3%; Pred. No. 1.3e-31;  
 Matches 213; Conservative 8; Mismatches 6; Indels 45; Gaps 1;

QY 2 DKTHTCPPCAPPELLGSPVFLFPKPKDTLMISRTPEVTCVVVDVSHEDPEVKFNWYVD 16  
 |||  
 Db 106 DTHTTCPPCAPPELLGSPVFLFPKPKDTLMISRTPEVTCVVVDVSHEDPEVKFNWYVD 165  
 |||

QY 17 GGPSVFLFPKPKDTLMISRTPEVTCVVVDVSHEDPEVKFNWYVDGVEVHNAKTKPREQ 76  
 |||  
 Db 166 GGPSVFLFPKPKDTLMISRTPEVTCVVVDVSHEDPEVKFNWYVDGVEVHNAKTKPREQ 225  
 |||  
 QY 77 YNSTYRVSVTLVHODWLNQKEYCKVSNKALPAPIETISKAKGQPREPQVYTLPPSR 136  
 |||  
 Db 226 YNSTYRVSVTLVHODWLNQKEYCKVSNKALPAPIETISKAKGQPREPQVYTLPPSR 285  
 |||  
 QY 137 DELTKNQVSLTCLVKGFYPSDIAVEWESNGQPENNYKTPPVLDSDGSFFLYSKLTVDKS 196  
 |||  
 Db 286 EEMTKNQVSLTCLVKGFYPSDIAVEWESNGQPENNYKTPPVLDSDGSFFLYSKLTVDKS 345  
 |||  
 QY 197 RWOQGNVFCSCVMHEALHNHYTKQSLSPGK 228  
 |||  
 Db 346 RWOQGNVFCSCVMHEALHNHYTKQSLSPGK 377  
 |||

## RESULT 6

A60764  
 Ig gamma-3 chain C region, form LAT - human  
 C:Species: Homo sapiens (man)  
 C:Date: 14-May-1993 #sequence\_revision 14-May-1993 #text\_change 16-Jul-1999  
 C:Accession: A60764  
 R: Huck, S.; Lefranc, G.; Lefranc, M.P.  
 Immunogenetics 30, 250-257, 1989  
 A:Title: A human immunoglobulin IGHG3 allele (Cmb0, b1, c3, c5, u) with an IGHG4 con  
 A:Reference number: A60764; MUID:90007613; PMID:25711587  
 A:Accession: A60764  
 A:Status: preliminary  
 A:Molecule type: DNA  
 A:Residues: 1-377 <HUC>  
 C:Superfamily: immunoglobulin C region; immunoglobulin homology  
 C:Keywords: immunoglobulin  
 F:20-85/Domain: immunoglobulin homology <IMM>

Query Match 86.7%; Score 1156; DB 2; Length 377;  
 Best Local Similarity 78.3%; Pred. No. 1.5e-31;  
 Matches 213; Conservative 8; Mismatches 6; Indels 45; Gaps 1;

QY 2 DKTHTCPPCAPPELLGSPVFLFPKPKDTLMISRTPEVTCVVVDVSHEDPEVKFNWYVDGVEVHNAKTKPREQ 16  
 |||  
 Db 106 DTHTTCPPCAPPELLGSPVFLFPKPKDTLMISRTPEVTCVVVDVSHEDPEVKFNWYVDGVEVHNAKTKPREQ 165  
 |||  
 QY 17 GGPSVFLFPKPKDTLMISRTPEVTCVVVDVSHEDPEVKFNWYVDGVEVHNAKTKPREQ 76  
 |||  
 Db 166 GGPSVFLFPKPKDTLMISRTPEVTCVVVDVSHEDPEVKFNWYVDGVEVHNAKTKPREQ 225  
 |||  
 QY 77 YNSTYRVSVTLVHODWLNQKEYCKVSNKALPAPIETISKAKGQPREPQVYTLPPSR 136  
 |||  
 Db 226 YNSTYRVSVTLVHODWLNQKEYCKVSNKALPAPIETISKAKGQPREPQVYTLPPSR 285  
 |||  
 QY 137 DELTKNQVSLTCLVKGFYPSDIAVEWESNGQPENNYKTPPVLDSDGSFFLYSKLTVDKS 196  
 |||  
 Db 286 EEMTKNQVSLTCLVKGFYPSDIAVEWESNGQPENNYKTPPVLDSDGSFFLYSKLTVDKS 345  
 |||  
 QY 197 RWOQGNVFCSCVMHEALHNHYTKQSLSPGK 228  
 |||  
 Db 346 RWOQGNVFCSCVMHEALHNHYTKQSLSPGK 377  
 |||

## RESULT 7

G2HU  
 Ig gamma-2 chain C region - human  
 C:Species: Homo sapiens (man)  
 C:Date: 30-Apr-1981 #sequence\_revision 13-Jun-1983 #text\_change 21-Jul-2000  
 C:Accession: A93906; A92809; A90752; A93132; A02148  
 R: Ellison, J.; Hood, L.  
 Proc. Natl. Acad. Sci. U.S.A. 79, 1984-1988, 1982  
 A:Title: Linkage and sequence homology of two human immunoglobulin gamma heavy chain  
 A:Reference number: A93906; MUID:82197621; PMID:6804948  
 A:Accession: A93906  
 A:Molecule type: DNA

A:Residues: 1-330 <ELL>  
A:Cross-references: EMBL:Z17370  
A:Note: This sequence has the G1m(17) allotypic marker, 97-Lys, and the G1m(1) markers,  
R: Haggis, L.J.  
submitted to the EMBL Data Library, October 1992  
A:Reference number: S33904  
A:Accession: S36861  
A:Molecule type: DNA  
A:Residues: 2-330 <HAR>  
A:Cross-references: EMBL:Z17370  
R:Takahashi, N.; Ueda, S.; Obata, M.; Nikaido, T.; Nakai, S.; Honjo, T.  
Cell, 29, 671-679, 1982  
A:Title: Structure of human immunoglobulin gamma genes: Implications for evolution of a  
A:Reference number: S33887; MUID:83001943; PMID:6811139  
A:Accession: S33887  
A:Molecule type: DNA  
A:Residues: 88-113; 235-330 <TAK>  
A:Cross-references: EMBL:Z17370  
R:Cunningham, B.A.; Rutishauser, U.; Gall, W.E.; Gottlieb, P.D.; Waxdal, M.J.; Edelman,  
Biochemistry 9, 3161-3170, 1970  
A:Title: The covalent structure of a human gammaG-immunoglobulin. VII. Amino acid sequen  
A:Reference number: A90563; MUID:71064024; PMID:5489771  
A:Contents: myeloma protein Eu  
A:Accession: B90563  
A:Molecule type: protein  
A:Residues: 1-96, 'R', 98-135 <CUN>  
A:Note: This sequence has the G1m(3) marker, 97-Arg  
R: Rutishauser, U.; Cunningham, B.A.; Bennett, C.; Konigsberg, W.H.; Edelman, G.M.  
Biochemistry 9, 3171-3181, 1970  
A:Title: The covalent structure of a human gammaG-immunoglobulin. VIII. Amino acid sequen  
A:Reference number: A90564; MUID:71064025; PMID:5530842  
A:Contents: Eu  
A:Accession: A90564  
A:Molecule type: protein  
A:Residues: 136-154, 'Q', 156-165, 'Q', 167-176, 'Q', 178-194, 'N', 196-197, 'D', 199-238, 'E', 240,  
A:Note: This sequence has the G1m(non-1) markers, 239-Glu and 241-Met  
R: Ponting, H.; Hilschmann, N.  
Hoppe-Seyler's Z. Physiol. Chem. 357, 1571-1604, 1976  
A:Title: Die Primärstruktur eines monoklonalen IgG1-Immunglobulins (Myelomprotein Nle)  
igen Primärstruktur.  
A:Reference number: A91668; MUID:77070269; PMID:826475  
A:Contents: myeloma protein Nle  
A:Accession: B91668  
A:Molecule type: protein  
A:Residues: 1-34, 'Q', 36-96, 'K', 98-115, 'Q', 117-197, 'D', 199-238, 'D', 240, 'L', 242-268, 'E', 27  
A:Note: This sequence has the G1m(17) and G1m(1) markers  
R: Schmidt, W.E.; Jung, H.D.; Palm, W.; Hilschmann, N.  
Hoppe-Seyler's Z. Physiol. Chem. 364, 713-747, 1983  
A:Title: Die Primärstruktur des kristallisierten monoklonalen Immunglobulins IgG1 KOL  
A:Reference number: A91723; MUID:83289131; PMID:6884994  
A:Contents: myeloma protein KOL; disulfide bonds  
A:Accession: A91723  
A:Molecule type: protein  
A:Residues: 1-96, 'R', 98-197, 'D', 199-238, 'E', 240, 'M', 242-266, 'D', 268-271, 'D', 273-330 <SCH  
A:Note: This sequence has the G1m(3) and G1m(non-1) markers  
R: Gall, W.E.; Edelman, G.M.  
Biochemistry 9, 3188-3196, 1970  
A:Title: The covalent structure of a human gammaG-immunoglobulin. X. Intrachain disulfid  
A:Reference number: A90565; MUID:71064027; PMID:4923144  
A:Contents: annotation; disulfide bonds  
R: Draker, L.; Schwarz, J.; Reichel, W.; Hilschmann, N.  
Hoppe-Seyler's Z. Physiol. Chem. 357, 1515-1540, 1976  
A:Title: Rule of antibody structure. The primary structure of monoclonal IgG1 immunoglob  
enbromide cleavage products, and the disulfide bridges.  
A:Reference number: A91667; MUID:77070267; PMID:1002129  
A:Contents: annotation; disulfide bonds  
A:Accession: B91667  
A:Gene: IGHG1  
A:Cross-references: GDB:120085; OMIM:147100  
A:Map position: 14q32.33-14q32.33  
A:Introns: 99/1; 114/1; 224/1  
C:Complex: An immunoglobulin heterotetramer subunit consists of two identical light (kap

hain disulfide bonds. In some cases, such as IgA and IgM, the subunits associate into  
C:Superfamily: immunoglobulin C region; immunoglobulin homology  
C:Keywords: duplication; glycoprotein; heterotetramer; immunoglobulin  
F:20-85/Domain: immunoglobulin homology <IM1>  
F:137-206/Domain: immunoglobulin homology <IM2>  
F:243-310/Domain: immunoglobulin homology <IM3>  
F:27-83,144-204,250-308/Disulfide bonds: #status experimental  
F:103/Disulfide bonds: interchain (to light chain) #status experimental  
F:109,112/Disulfide bonds: interchain (to heavy chain) #status experimental  
F:180/Blinding site: carbohydrate (Asn) (covalent) #status experimental

Query Match 92.48; Score 1233; DB 1; Length 330;  
Best Local Similarity 100.0%; Pred. No. 2.3e-34;  
Matches 227; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2 DKTHCPCPAPELLGGPSVFLPFPKPKDTLMISRTPEVTCVVVDVSHEDPEVKFNWYVD 61

Db 104 DKTHCPCPAPELLGGPSVFLPFPKPKDTLMISRTPEVTCVVVDVSHEDPEVKFNWYVD 163

Qy 62 GVEVHNAKTKPREQYNSTYRVVSVLTVLHQDWLNGKEYCKVSNKALPAPIEKTISKAK 121

Db 164 GVEVHNAKTKPREQYNSTYRVVSVLTVLHQDWLNGKEYCKVSNKALPAPIEKTISKAK 223

Qy 122 GQPREPQVYTLPPSRDELTKNQVSLTCLVKGFYPSDIAVEWESNGQPENNYKTTTPVLD 181

Db 224 GQPREPQVYTLPPSRDELTKNQVSLTCLVKGFYPSDIAVEWESNGQPENNYKTTTPVLD 283

Qy 182 DGSFFLYSKLVDRKRWQGVNFCVSMHEALHNYTKSLSPGK 228

Db 284 DGSFFLYSKLVDRKRWQGVNFCVSMHEALHNYTKSLSPGK 330

# RESULT 3

S69339

Ig heavy chain V region precursor - human

C:Species: Homo sapiens (man)

C:Date: 19-Mar-1997 #sequence\_revision 19-Mar-1997 #text\_change 01-Dec-2000

C:Accession: S69339; S72664

R:Khamilichi, A.A.; Aucouturier, P.; Preud'homme, J.L.; Cogne, M.

Eur. J. Biochem. 229, 54-60, 1995

A:Title: Structure of abnormal heavy chains in human heavy-chain-deposition disease.

A:Reference number: S69339; MUID:95262687; PMID:7744049

A:Accession: S69339

A>Status: preliminary

A:Molecule type: mRNA

A:Residues: 1-374 <KHA>

A:Cross-references: EMBL:X81695

R:Khamilichi, A.A.

submitted to the EMBL Data Library, September 1994

A:Reference number: S72664

A:Accession: S72664

A>Status: preliminary

A:Molecule type: mRNA

A:Residues: 1-140, 'C', 142-374 <KH2>

A:Cross-references: EMBL:X81695

C:Superfamily: immunoglobulin C region; immunoglobulin homology

Query Match 92.08; Score 1227; DB 2; Length 374;

Best Local Similarity 99.18; Pred. No. 5.5e-34;

Matches 225; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

Qy 2 DKTHCPCPAPELLGGPSVFLPFPKPKDTLMISRTPEVTCVVVDVSHEDPEVKFNWYVD 61

Db 148 DKTHCPCPAPELLGGPSVFLPFPKPKDTLMISRTPEVTCVVVDVSHEDPEVKFNWYVD 207

Qy 62 GVEVHNAKTKPREQYNSTYRVVSVLTVLHQDWLNGKEYCKVSNKALPAPIEKTISKAK 121

Db 208 GVEVHNAKTKPREQYNSTYRVVSVLTVLHQDWLNGKEYCKVSNKALPAPIEKTISKAK 267

Qy 122 GQPREPQVYTLPPSRDELTKNQVSLTCLVKGFYPSDIAVEWESNGQPENNYKTTTPVLD 181

Db 268 GQPREPQVYTLPPSRDELTKNQVSLTCLVKGFYPSDIAVEWESNGQPENNYKTTTPVLD 327

GenCore version 5.1.4.p5\_4578  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: April 21, 2003, 10:42:26 ; Search time 18.0882 Seconds  
(without alignments)  
1307.428 Million cell updates/sec

Title: 2LINK7  
Perfect score: 1334  
Sequence: 1 MDKTHTCPPCPAPPELLGGPS.....GKDWLKAFYDKVAEKLKEAF 246

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.0

Searched: 283224 seqs, 96134422 residues  
Total number of hits satisfying chosen parameters: 283224

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000  
Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : PIR\_73:.\*  
1: pir1:.\*  
2: pir2:.\*  
3: pir3:.\*  
4: pir4:.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1233	92.4	255	4	S31866
2	1233	92.4	330	1	GHU
3	1227	92.0	374	2	S69339
4	1180	88.5	234	2	PT0207
5	1158	86.8	377	2	A23511
6	1156	86.7	377	2	A60764
7	1146	85.9	326	1	G2HU
8	1142	85.6	327	1	G4HU
9	1133	84.9	289	1	G3HUI
10	921	69.0	323	1	GHRB
11	917	68.7	328	2	I47160
12	917	68.7	328	2	I47159
13	911	68.3	277	2	I47162
14	894	67.0	329	1	G2GP
15	888	66.6	328	2	I47158
16	885	66.3	328	2	I47161
17	860	64.5	470	2	S22080
18	858	64.3	472	2	S31459
19	853	63.9	308	2	C30554
20	848	63.6	329	1	G3MSC
21	840	63.0	333	2	PS0018
22	838	62.8	398	1	G3MSM
23	834	62.5	444	2	PC4436
24	825	61.8	326	2	PS0017
25	823	61.7	324	1	G1MS
26	823	61.7	393	1	G1MSM
27	819	61.4	399	1	G2MSAM
28	813	61.3	329	2	S00847
29	813	60.9	330	1	G2MSA

30	813	60.9	469	2	S37483
31	809	60.6	335	1	G2MSAB
32	799	59.9	446	2	S40295
33	793	59.4	322	2	PS0019
34	791	59.3	405	1	G2MSBM
35	787	59.0	474	1	G2MS11
36	774	58.0	327	2	S06611
37	766	57.4	475	2	S01321
38	707	53.0	180	2	I46732
39	602	45.1	218	2	A36040
40	601	45.1	249	2	S69340
41	571	42.8	152	2	S14236
42	466	34.9	26926	1	I38344
43	437	32.8	572	2	B46529
44	422	31.6	548	2	S38864
45	407	30.5	448	2	S03186

ALIGNMENTS

RESULT 1

S31866

Ig gamma-1 chain C region - synthetic

C:Species: synthetic

A:Note: Homo sapiens (man) gene engineered and expressed in Escherichia coli

C:Date: 06-Jan-1995 #sequence\_revision 17-Mar-1997 #text\_change 19-May-2000

C:Accession: S31866

R:Filipula, D.

submitted to the EMBL Data Library, February 1993

A:Description: Screening method for protein-protein interactions of cloned gene products

A:Reference number: S31866

A:Accession: S31866

A:Molecule type: mRNA

A:Residues: 1-255 <FILL>

A:Cross-references: EMBL:X70421; NID:g33068; PIDN:CAA49866.1; PID:g33069

C:Keywords: immunoglobulin

F:1-22/Region: Escherichia coli outer membrane protein A precursor

F:23-255/Region: human Ig gamma-1 chain C region

Query Match

Best Local Similarity 92.4%; Score 1233; DB 4; Length 255;

Matches 227; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 DKTHTCPPCPAPPELLGGPSVFLPSPKDTLMISRTPEVTCVVVDVSHEDPEVKFNWYD 61

DB 29 DKTHTCPPCPAPPELLGGPSVFLPSPKDTLMISRTPEVTCVVVDVSHEDPEVKFNWYD 88

QY 62 GVEVHNAKTPREEQYNSTYRVVSVLTVLHQDWLNGKEYKCKVSNKALPAPIEKTISKAK 121

DB 89 GVEVHNAKTPREEQYNSTYRVVSVLTVLHQDWLNGKEYKCKVSNKALPAPIEKTISKAK 148

QY 122 GQPRFPQVYTLPPSRDELTKNQVSLTCLVKGPYPSDIAVEWESNGQPENNYKTPPVLD 181

DB 149 GQPRFPQVYTLPPSRDELTKNQVSLTCLVKGPYPSDIAVEWESNGQPENNYKTPPVLD 208

QY 182 DGSFFLYSKLTVDKSRWQGNVFSCVMEALHNHYTQKSLSLSPGK 228

DB 209 DGSFFLYSKLTVDKSRWQGNVFSCVMEALHNHYTQKSLSLSPGK 255

RESULT 2

GHU

Ig gamma-1 chain C region - human

C:Species: Homo sapiens (man)

C:Date: 31-Jan-1981 #sequence\_revision 18-Aug-1982 #text\_change 16-Jul-1999

C:Accession: A93433; S36861; S33887; B90563; A90564; B91668; A91723; A02146

R:Ellison, J.W.; Berson, B.J.; Hood, L.E.

Nucleic Acids Res. 10, 4071-4079, 1982

A:Title: The nucleotide sequence of a human immunoglobulin C-gamma gene.

A:Reference number: A93433; MUID:82274238; PMID:6287432

A:Accession: A93433

A:Molecule type: DNA





Matches	227;	Conservative	1;	Mismatches	0;	Indels	0;	Gaps	0;
QY	1	MDKTHTCPPCPAPPELLGGPSVFLFPKPKDTLMISRTPEVTCVVVDVSHEDPEVKFNWYV	60						
Db	2	VDKTHTCPPCPAPPELLGGPSVFLFPKPKDTLMISRTPEVTCVVVDVSHEDPEVKFNWYV	61						
QY	61	DGVEVHNAKTKPREEQYNSTYRVVSVLTVLHQDWLNGKEYKCKVSNKALPAPIEKTISKA	120						
Db	62	DGVEVHNAKTKPREEQYNSTYRVVSVLTVLHQDWLNGKEYKCKVSNKALPAPIEKTISKA	121						
QY	121	KGQPREPVYTLPPSRDELTKNQVSLTCLVKGYFSPDIAVEWESNGQPENNYKTTTPPVL	180						
Db	122	KGQPREPVYTLPPSRDELTKNQVSLTCLVKGYFSPDIAVEWESNGQPENNYKTTTPPVL	181						
QY	181	SDGSFFLYSKLTVDKSRWQOGNVFSCVMHEALHNHYTQKSLSLSPGK	228						
Db	182	SDGSFFLYSKLTVDKSRWQOGNVFSCVMHEALHNHYTQKSLSLSPGK	229						

Search completed: April 21, 2003, 10:51:15  
Job time : 13.3353 secs

```

; GENERAL INFORMATION:
; APPLICANT: Young, Robert
; TITLE OF INVENTION: Compounds for Targeting
; FILE REFERENCE: 43191-256808
; CURRENT APPLICATION NUMBER: US/09/825,012
; CURRENT FILING DATE: 2001-04-03
; PRIOR APPLICATION NUMBER: US 60/237,159
; PRIOR FILING DATE: 2000-10-02
; PRIOR APPLICATION NUMBER: GB 0008049.9
; PRIOR FILING DATE: 2000-04-03
; NUMBER OF SEQ ID NOS: 102
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 49
; LENGTH: 730
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Humanised HMFg1 heavy chain - DNase I fusion
US-09-825-012-49

Query Match          92.7%; Score 1237; DB 10; Length 730;
Best Local Similarity 48.7%; Pred. No. 6.4e-32;
Matches 233; Conservative 6; Mismatches 5; Indels 234; Gaps 4;

QY  2 DKHTCCPPCAPPELLGGPSVFLPPKPKDTLMISRTPEVTCVVVDVSHEDPEVKFNWYVD 61
Db  241 DKHTCCPPCAPPELLGGPSVFLPPKPKDTLMISRTPEVTCVVVDVSHEDPEVKFNWYVD 300

QY  62 GVEVHNAKTPREEQNSYRYVSVLTVLHQDLNKGKEYCKVSNKALPAPIETISKAK 121
Db  301 GVEVHNAKTPREEQNSYRYVSVLTVLHQDLNKGKEYCKVSNKALPAPIETISKAK 360

QY  122 GQPREPQVYTLPPSRDELTKNQVSLTCLVKGFYPSDIAVWESNGQPENNYKTTPPVLD 181
Db  361 GQPREPQVYTLPPSRDELTKNQVSLTCLVKGFYPSDIAVWESNGQPENNYKTTPPVLD 420

QY  182 DGSFFLYSKLTVDKSRWQQGNVSCVMHEALHNNHYTKQSLSLSP----- 226
Db  421 DGSFFLYSKLTVDKSRWQQGNVSCVMHEALHNNHYTKQSLSLSP----- 480

QY  227 ----- 226
Db  481 FGETKSNATLVSIVQILSRDYIALVQEVDRSHLTAVGKLLDNLNQADPDYHYVVS 540

QY  227 -----GKD----- 229
Db  541 LGRNSYKERYLFYRPDQVSAVDYDDGCEPCGNDTENREPAIVRFFSRTTEVREFAI 600

QY  230 ----- 229
Db  601 VPLHAAPGDAVAEIDALDYLDVQEKWGLEDMYLMGDFNAGCSYVRPSQWSSIRLWTS 660

QY  230 ---WL-----KAFYDKV-----AEKLEA 245
Db  661 TFQWLIPDSADTTATTHCAIDRVVAGMLLRGAVVPSALPFPFOAAYGLSDQLAQA 718

RESULT 14
US-09-825-012-58
; Sequence 58, Application US/09825012
; Patent No. US2002012798A1
; GENERAL INFORMATION:
; APPLICANT: Young, Robert
; TITLE OF INVENTION: Compounds for Targeting
; FILE REFERENCE: 43191-256808
; CURRENT APPLICATION NUMBER: US/09/825,012
; CURRENT FILING DATE: 2001-04-03
; PRIOR APPLICATION NUMBER: US 60/237,159
; PRIOR FILING DATE: 2000-10-02
; PRIOR APPLICATION NUMBER: GB 0008049.9
; PRIOR FILING DATE: 2000-04-03
; NUMBER OF SEQ ID NOS: 102
; SOFTWARE: PatentIn version 3.1

```

```

; SEQ ID NO 58
; LENGTH: 740
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Humanised HMFg1 heavy chain - DNase I fusion
US-09-825-012-58

Query Match          92.7%; Score 1237; DB 10; Length 740;
Best Local Similarity 48.7%; Pred. No. 6.6e-32;
Matches 233; Conservative 6; Mismatches 5; Indels 234; Gaps 4;

QY  2 DKHTCCPPCAPPELLGGPSVFLPPKPKDTLMISRTPEVTCVVVDVSHEDPEVKFNWYVD 61
Db  241 DKHTCCPPCAPPELLGGPSVFLPPKPKDTLMISRTPEVTCVVVDVSHEDPEVKFNWYVD 300

QY  62 GVEVHNAKTPREEQNSYRYVSVLTVLHQDLNKGKEYCKVSNKALPAPIETISKAK 121
Db  301 GVEVHNAKTPREEQNSYRYVSVLTVLHQDLNKGKEYCKVSNKALPAPIETISKAK 360

QY  122 GQPREPQVYTLPPSRDELTKNQVSLTCLVKGFYPSDIAVWESNGQPENNYKTTPPVLD 181
Db  361 GQPREPQVYTLPPSRDELTKNQVSLTCLVKGFYPSDIAVWESNGQPENNYKTTPPVLD 420

QY  182 DGSFFLYSKLTVDKSRWQQGNVSCVMHEALHNNHYTKQSLSLSP----- 226
Db  421 DGSFFLYSKLTVDKSRWQQGNVSCVMHEALHNNHYTKQSLSLSP----- 480

QY  227 ----- 226
Db  481 FGETKSNATLVSIVQILSRDYIALVQEVDRSHLTAVGKLLDNLNQADPDYHYVVS 540

QY  227 -----GKD----- 229
Db  541 LGRNSYKERYLFYRPDQVSAVDYDDGCEPCGNDTENREPAIVRFFSRTTEVREFAI 600

QY  230 ----- 229
Db  601 VPLHAAPGDAVAEIDALDYLDVQEKWGLEDMYLMGDFNAGCSYVRPSQWSSIRLWTS 660

QY  230 ---WL-----KAFYDKV-----AEKLEA 245
Db  661 TFQWLIPDSADTTATTHCAIDRVVAGMLLRGAVVPSALPFPFOAAYGLSDQLAQA 718

RESULT 15
US-10-215-297-2
; Sequence 2, Application US/10215297
; Publication No. US2002019222A1
; GENERAL INFORMATION:
; APPLICANT: Blumberg, Richard S.
; APPLICANT: Simister, Neil E.
; APPLICANT: Lencer, Wayne I.
; TITLE OF INVENTION: RECEPTOR SPECIFIC TRANSEPITHELIAL TRANSPORT OF
; TITLE OF INVENTION: THERAPEUTICS
; FILE REFERENCE: S1383/77003
; CURRENT APPLICATION NUMBER: US/10/215,297
; CURRENT FILING DATE: 2002-08-08
; PRIOR APPLICATION NUMBER: US 08/578,171
; PRIOR FILING DATE: 1995-12-29
; PRIOR APPLICATION NUMBER: US 08/374,159
; PRIOR FILING DATE: 1995-01-17
; PRIOR APPLICATION NUMBER: US 09/122,144
; PRIOR FILING DATE: 1998-07-24
; NUMBER OF SEQ ID NOS: 22
; SEQ ID NO 2
; LENGTH: 229
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-215-297-2

Query Match          92.5%; Score 1234; DB 9; Length 229;
Best Local Similarity 99.6%; Pred. No. 5.5e-33;

```

;; PRIOR FILING DATE: 2000-05-03  
;; NUMBER OF SEQ ID NOS: 135  
;; SOFTWARE: PatentIn version 3.1  
;; SEQ ID NO 109  
;; LENGTH: 282  
;; TYPE: PRT  
;; ORGANISM: Artificial Sequence  
;; OTHER INFORMATION: Echistatin Fc-peptide  
;; NAME/KEY: misc\_feature  
;; LOCATION: (1)..(1)  
;; OTHER INFORMATION: NdeI site  
;; NAME/KEY: misc\_feature  
;; LOCATION: (854)..(854)  
;; OTHER INFORMATION: BamHI site  
US-09-840-277-109

Query Match 92.8%; Score 1238; DB 9; Length 282;  
Best Local Similarity 100.0%; Pred. No. 6.e-33;  
Matches 228; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MDKTHTCPPCAPPELLGGPSVFLPPKPKDTLMISRTPEVTCVVDVSHEDPEVKFNWYV 60  
Db 1 MDKTHTCPPCAPPELLGGPSVFLPPKPKDTLMISRTPEVTCVVDVSHEDPEVKFNWYV 60  
QY 61 DGEVHNAKTPREEQYNSTYRVVSVLTVLHQDWLNGKEYCKVSNKALPAPIEKTISK 120  
Db 61 DGEVHNAKTPREEQYNSTYRVVSVLTVLHQDWLNGKEYCKVSNKALPAPIEKTISK 120  
QY 121 KGPREQVYTLPPSRDELTKNQVSLTCLVKGFYPSDIAVEWESNGQPENNYKTPPVLD 180  
Db 121 KGPREQVYTLPPSRDELTKNQVSLTCLVKGFYPSDIAVEWESNGQPENNYKTPPVLD 180  
QY 181 SDGSFFLYSKLTVDKSRWQGNVFSCSVMHHEALHNHYTQKSLSLSPGK 228  
Db 181 SDGSFFLYSKLTVDKSRWQGNVFSCSVMHHEALHNHYTQKSLSLSPGK 228

RESULT 11  
US-09-784-623-14  
;; Sequence 14, Application US/09784623  
;; Patent No. US200200094541  
;; GENERAL INFORMATION:  
;; APPLICANT: Boone, Thomas C.  
;; APPLICANT: Hershenson, Susan  
;; APPLICANT: Bevilacqua, Michael P.  
;; APPLICANT: Collins, David S.  
;; TITLE OF INVENTION: COMPOSITION AND METHOD FOR TREATING INFLAMMATORY  
;; DISEASES  
;; FILE REFERENCE: A-365F  
;; CURRENT APPLICATION NUMBER: US/09/784,623  
;; CURRENT FILING DATE: 2001-02-15  
;; PRIOR APPLICATION NUMBER: 09/131,247  
;; PRIOR FILING DATE: 1998-08-07  
;; PRIOR APPLICATION NUMBER: PCT/US 97/02131  
;; PRIOR FILING DATE: 1997-02-10  
;; NUMBER OF SEQ ID NOS: 16  
;; SOFTWARE: PatentIn ver. 2.0  
;; SEQ ID NO 14  
;; LENGTH: 389  
;; TYPE: PRT  
;; ORGANISM: Human  
US-09-784-623-14

Query Match 92.8%; Score 1238; DB 10; Length 389;  
Best Local Similarity 72.7%; Pred. No. 1.4e-32;  
Matches 234; Conservative 4; Mismatches 5; Indels 79; Gaps 4;

QY 2 DKHTCPCCAPPELLGGPSVFLPPKPKDTLMISRTPEVTCVVDVSHEDPEVKFNWYV 61  
Db 10 DKHTCPCCAPPELLGGPSVFLPPKPKDTLMISRTPEVTCVVDVSHEDPEVKFNWYV 69  
QY 62 GVEVHNAKTPREEQYNSTYRVVSVLTVLHQDWLNGKEYCKVSNKALPAPIEKTISK 121

Db 70 GVEVHNAKTPREEQYNSTYRVVSVLTVLHQDWLNGKEYCKVSNKALPAPIEKTISK 129  
QY 122 GQPREQVYTLPPSRDELTKNQVSLTCLVKGFYPSDIAVEWESNGQPENNYKTPPVLD 181  
Db 130 GQPREQVYTLPPSRDELTKNQVSLTCLVKGFYPSDIAVEWESNGQPENNYKTPPVLD 189  
QY 182 DGSFFLYSKLTVDKSRWQGNVFSCSVMHHEALHNHYTQKSLSLSPGK----- 228  
Db 190 DGSFFLYSKLTVDKSRWQGNVFSCSVMHHEALHNHYTQKSLSLSPGKMKRPSGRKSKMQA 249  
QY 229 --DW---LKAFY-----DKV----- 238  
Db 250 FRWDVNAQKTFYLRNNOLVAGYLOGPNVLEEKIDVVPPIEPHALFLGIHGKMLSCVKS 309  
QY 239 -----AEKLE 244

Db 310 GDETRLOLEAVNITDLSNRKQ 331  
RESULT 12  
US-09-854-864-31  
;; Sequence 31, Application US/09854864  
;; Patent No. US20020081296A1  
;; GENERAL INFORMATION:  
;; APPLICANT: THEILL, LARS EYDE  
;; APPLICANT: YU, GANG  
;; TITLE OF INVENTION: METHODS AND COMPOSITIONS OF MATTER CONCERNING APRIL/G70, BCMA  
;; FILE REFERENCE: B15/AGP-3, AND TACI  
;; CURRENT APPLICATION NUMBER: US/09/854,864  
;; CURRENT FILING DATE: 2001-09-11  
;; PRIOR APPLICATION NUMBER: US 60/204,039  
;; PRIOR FILING DATE: 2000-05-12  
;; PRIOR APPLICATION NUMBER: US 60/214,591  
;; PRIOR FILING DATE: 2000-06-27  
;; NUMBER OF SEQ ID NOS: 31  
;; SOFTWARE: PatentIn version 3.1  
;; SEQ ID NO 31  
;; LENGTH: 394  
;; TYPE: PRT  
;; ORGANISM: Homo sapiens  
US-09-854-864-31

Query Match 92.7%; Score 1237; DB 10; Length 394;  
Best Local Similarity 78.6%; Pred. No. 1.5e-32;  
Matches 231; Conservative 1; Mismatches 5; Indels 57; Gaps 2;

QY 2 DKHTCPCCAPPELLGGPSVFLPPKPKDTLMISRTPEVTCVVDVSHEDPEVKFNWYV 61  
Db 20 DKHTCPCCAPPELLGGPSVFLPPKPKDTLMISRTPEVTCVVDVSHEDPEVKFNWYV 79  
QY 62 GVEVHNAKTPREEQYNSTYRVVSVLTVLHQDWLNGKEYCKVSNKALPAPIEKTISK 121  
Db 80 GVEVHNAKTPREEQYNSTYRVVSVLTVLHQDWLNGKEYCKVSNKALPAPIEKTISK 139  
QY 122 GQPREQVYTLPPSRDELTKNQVSLTCLVKGFYPSDIAVEWESNGQPENNYKTPPVLD 181  
Db 140 GQPREQVYTLPPSRDELTKNQVSLTCLVKGFYPSDIAVEWESNGQPENNYKTPPVLD 199  
QY 182 DGSFFLYSKLTVDKSRWQGNVFSCSVMHHEALHNHYTQKSLSLSPGK----- 229  
Db 200 DGSFFLYSKLTVDKSRWQGNVFSCSVMHHEALHNHYTQKSLSLSPGKSRAVLTKOKKQH 259  
QY 230 -----WLKA-----FYDKV 238  
Db 260 SVLHLVPINATSKDDSDVTEVMQPALRRGRGLQAQGYGVRIQDAGVYLLYSQV 313

RESULT 13  
US-09-825-012-49  
;; Sequence 49, Application US/09825012  
;; Patent No. US2002012798A1

```
; Sequence 2, Application US/09840669B
; Publication No. US2003004070A1
; GENERAL INFORMATION:
; APPLICANT: KOHNO, TADAHIKO
; TITLE OF INVENTION: APO-AI/II PEPTIDE DERIVATIVES
; FILE REFERENCE: A-690
; CURRENT APPLICATION NUMBER: US/09/840,669B
; CURRENT FILING DATE: 2002-06-07
; PRIOR APPLICATION NUMBER: 60/198,920
; PRIOR FILING DATE: 2000-04-21
; NUMBER OF SEQ ID NOS: 11
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 2
; LENGTH: 228
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-840-669B-2

Query Match          92.8%; Score 1238; DB 9; Length 228;
Best Local Similarity 100.0%; Pred. No. 4e-33;
Matches 228; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MDKTHCTPCPAPELLGGPSVFLFPPPKDPTLMISRTPEVTCVVVDVSHEDPEVKFNWYV 60
DB 1 MDKTHCTPCPAPELLGGPSVFLFPPPKDPTLMISRTPEVTCVVVDVSHEDPEVKFNWYV 60
QY 61 DGVEVHNKTKPREEQYNSTYRVSVSLTVLHODWLNKKEYCKVSKNKPAPIEKTISKA 120
DB 61 DGVEVHNKTKPREEQYNSTYRVSVSLTVLHODWLNKKEYCKVSKNKPAPIEKTISKA 120
QY 121 KGQPREPQVYTLPPSRDELTKNOVSLTCLVKGFPYPSDIAVEWESNGQPENNYKTTTPPVLD 180
DB 121 KGQPREPQVYTLPPSRDELTKNOVSLTCLVKGFPYPSDIAVEWESNGQPENNYKTTTPPVLD 180
QY 181 SDGSFFLYSKLTVDKSRWQQGNVFCSCVMHEALHNHYTOKSLSLSPGK 228
DB 181 SDGSFFLYSKLTVDKSRWQQGNVFCSCVMHEALHNHYTOKSLSLSPGK 228

; Sequence 2, Application US/09843221A
; Publication No. US20030039654A1
; GENERAL INFORMATION:
; APPLICANT: KOSTENUIK, PAUL
; APPLICANT: LIU, CHUAN-FA
; APPLICANT: LACEY, DAVID LEE
; TITLE OF INVENTION: MODULATORS OF RECEPTORS FOR PARATHYROID HORMONE AND PARATHYROID H
; TITLE OF INVENTION: RELATED PROTEIN
; FILE REFERENCE: A-665B
; CURRENT APPLICATION NUMBER: US/09/843,221A
; CURRENT FILING DATE: 2001-04-26
; PRIOR APPLICATION NUMBER: 60/266,673
; PRIOR FILING DATE: 2001-02-06
; PRIOR APPLICATION NUMBER: 60/214,860
; PRIOR FILING DATE: 2000-06-28
; PRIOR APPLICATION NUMBER: 60/200,053
; PRIOR FILING DATE: 2000-04-27
; NUMBER OF SEQ ID NOS: 170
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 2
; LENGTH: 228
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-843-221A-2

Query Match          92.8%; Score 1238; DB 9; Length 228;
Best Local Similarity 100.0%; Pred. No. 4e-33;
Matches 228; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MDKTHCTPCPAPELLGGPSVFLFPPPKDPTLMISRTPEVTCVVVDVSHEDPEVKFNWYV 60
DB 1 MDKTHCTPCPAPELLGGPSVFLFPPPKDPTLMISRTPEVTCVVVDVSHEDPEVKFNWYV 60
QY 61 DGVEVHNKTKPREEQYNSTYRVSVSLTVLHODWLNKKEYCKVSKNKPAPIEKTISKA 120
DB 61 DGVEVHNKTKPREEQYNSTYRVSVSLTVLHODWLNKKEYCKVSKNKPAPIEKTISKA 120
QY 121 KGQPREPQVYTLPPSRDELTKNOVSLTCLVKGFPYPSDIAVEWESNGQPENNYKTTTPPVLD 180
DB 121 KGQPREPQVYTLPPSRDELTKNOVSLTCLVKGFPYPSDIAVEWESNGQPENNYKTTTPPVLD 180
QY 181 SDGSFFLYSKLTVDKSRWQQGNVFCSCVMHEALHNHYTOKSLSLSPGK 228
DB 181 SDGSFFLYSKLTVDKSRWQQGNVFCSCVMHEALHNHYTOKSLSLSPGK 228

; Sequence 2, Application US/09843221A
; Publication No. US20030039654A1
; GENERAL INFORMATION:
; APPLICANT: KOSTENUIK, PAUL
; APPLICANT: LIU, CHUAN-FA
; APPLICANT: LACEY, DAVID LEE
; TITLE OF INVENTION: MODULATORS OF RECEPTORS FOR PARATHYROID HORMONE AND PARATHYROID H
; TITLE OF INVENTION: RELATED PROTEIN
; FILE REFERENCE: A-665B
; CURRENT APPLICATION NUMBER: US/09/843,221A
; CURRENT FILING DATE: 2001-04-26
; PRIOR APPLICATION NUMBER: 60/266,673
; PRIOR FILING DATE: 2001-02-06
; PRIOR APPLICATION NUMBER: 60/214,860
; PRIOR FILING DATE: 2000-06-28
; PRIOR APPLICATION NUMBER: 60/200,053
; PRIOR FILING DATE: 2000-04-27
; NUMBER OF SEQ ID NOS: 170
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 2
; LENGTH: 228
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-843-221A-2

Query Match          92.8%; Score 1238; DB 9; Length 228;
Best Local Similarity 100.0%; Pred. No. 4e-33;
Matches 228; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MDKTHCTPCPAPELLGGPSVFLFPPPKDPTLMISRTPEVTCVVVDVSHEDPEVKFNWYV 60
DB 1 MDKTHCTPCPAPELLGGPSVFLFPPPKDPTLMISRTPEVTCVVVDVSHEDPEVKFNWYV 60
QY 61 DGVEVHNKTKPREEQYNSTYRVSVSLTVLHODWLNKKEYCKVSKNKPAPIEKTISKA 120
DB 61 DGVEVHNKTKPREEQYNSTYRVSVSLTVLHODWLNKKEYCKVSKNKPAPIEKTISKA 120
QY 121 KGQPREPQVYTLPPSRDELTKNOVSLTCLVKGFPYPSDIAVEWESNGQPENNYKTTTPPVLD 180
DB 121 KGQPREPQVYTLPPSRDELTKNOVSLTCLVKGFPYPSDIAVEWESNGQPENNYKTTTPPVLD 180
QY 181 SDGSFFLYSKLTVDKSRWQQGNVFCSCVMHEALHNHYTOKSLSLSPGK 228
DB 181 SDGSFFLYSKLTVDKSRWQQGNVFCSCVMHEALHNHYTOKSLSLSPGK 228
```

```
QY 61 DGVEVHNKTKPREEQYNSTYRVSVSLTVLHODWLNKKEYCKVSKNKPAPIEKTISKA 120
DB 61 DGVEVHNKTKPREEQYNSTYRVSVSLTVLHODWLNKKEYCKVSKNKPAPIEKTISKA 120
QY 121 KGQPREPQVYTLPPSRDELTKNOVSLTCLVKGFPYPSDIAVEWESNGQPENNYKTTTPPVLD 180
DB 121 KGQPREPQVYTLPPSRDELTKNOVSLTCLVKGFPYPSDIAVEWESNGQPENNYKTTTPPVLD 180
QY 181 SDGSFFLYSKLTVDKSRWQQGNVFCSCVMHEALHNHYTOKSLSLSPGK 228
DB 181 SDGSFFLYSKLTVDKSRWQQGNVFCSCVMHEALHNHYTOKSLSLSPGK 228
```

## RESULT 9

```
US-09-847-712-2
; Sequence 2, Application US/09847712
; Patent No. US20020090646A1
; GENERAL INFORMATION:
; APPLICANT: LIU, CHUAN-FA
; APPLICANT: MARSHALL, WILLIAM S.
; APPLICANT: REYNOLDS, ANGELA
; TITLE OF INVENTION: CALCITONIN-RELATED MOLECULES
; FILE REFERENCE: A-684
; CURRENT APPLICATION NUMBER: US/09/847,712
; CURRENT FILING DATE: 2001-05-02
; PRIOR APPLICATION NUMBER: 60/201,511
; PRIOR FILING DATE: 2000-05-03
; NUMBER OF SEQ ID NOS: 26
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 2
; LENGTH: 228
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-847-712-2
```

```
Query Match          92.8%; Score 1238; DB 10; Length 228;
Best Local Similarity 100.0%; Pred. No. 4e-33;
Matches 228; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY 1 MDKTHCTPCPAPELLGGPSVFLFPPPKDPTLMISRTPEVTCVVVDVSHEDPEVKFNWYV 60
DB 1 MDKTHCTPCPAPELLGGPSVFLFPPPKDPTLMISRTPEVTCVVVDVSHEDPEVKFNWYV 60
QY 61 DGVEVHNKTKPREEQYNSTYRVSVSLTVLHODWLNKKEYCKVSKNKPAPIEKTISKA 120
DB 61 DGVEVHNKTKPREEQYNSTYRVSVSLTVLHODWLNKKEYCKVSKNKPAPIEKTISKA 120
QY 121 KGQPREPQVYTLPPSRDELTKNOVSLTCLVKGFPYPSDIAVEWESNGQPENNYKTTTPPVLD 180
DB 121 KGQPREPQVYTLPPSRDELTKNOVSLTCLVKGFPYPSDIAVEWESNGQPENNYKTTTPPVLD 180
QY 181 SDGSFFLYSKLTVDKSRWQQGNVFCSCVMHEALHNHYTOKSLSLSPGK 228
DB 181 SDGSFFLYSKLTVDKSRWQQGNVFCSCVMHEALHNHYTOKSLSLSPGK 228
```

## RESULT 10

```
US-09-840-277-109
; Sequence 109, Application US/09840277
; Patent No. US20020168363A1
; GENERAL INFORMATION:
; APPLICANT: FEIGE, ULRICH
; APPLICANT: KOHNO, TADAHIKO
; APPLICANT: LACEY, DAVID LEE
; APPLICANT: BOONE, THOMAS CHARLES
; TITLE OF INVENTION: INTEGRIN/ADHESION ANTAGONISTS
; FILE REFERENCE: A-688A
; CURRENT APPLICATION NUMBER: US/09/840,277
; CURRENT FILING DATE: 2001-08-14
; PRIOR APPLICATION NUMBER: 60/198,919
; PRIOR FILING DATE: 2000-04-21
; PRIOR APPLICATION NUMBER: 60/201,394
```

```
; APPLICANT: Young, Robert
; TITLE OF INVENTION: Compounds for Targeting
; FILE REFERENCE: 43191-256808
; CURRENT APPLICATION NUMBER: US/09/825,012
; CURRENT FILING DATE: 2001-04-03
; PRIOR APPLICATION NUMBER: US 60/237,159
; PRIOR FILING DATE: 2000-10-02
; PRIOR APPLICATION NUMBER: GB 0008049.9
; PRIOR FILING DATE: 2000-04-03
; NUMBER OF SEQ ID NOS: 102
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 61
; LENGTH: 739
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Humanised HMFG1 heavy chain - DNase I fusion
US-09-825-012-61

Query Match          93.3%; Score 1245; DB 10; Length 739;
Best Local Similarity 49.1%; Pred. No. 3.6e-32;
Matches 234; Conservative 6; Mismatches 4; Indels 233; Gaps 3;

Qy  2 DKTHCTPCPAPPELLGGPSVFLFPPKPKDTLMISRTPEVTCVVVDVSHEDPEVKFNMYD 61
    |||
Db  241 DKTHCTPCPAPPELLGGPSVFLFPPKPKDTLMISRTPEVTCVVVDVSHEDPEVKFNMYD 300
    |||

Qy  62 GVEVHNAKTRPREQYNSTYRVVSVLTVLHQDWLNGKEYKCKVSNKALPAPIEKTISKAK 121
    |||
Db  301 GVEVHNAKTRPREQYNSTYRVVSVLTVLHQDWLNGKEYKCKVSNKALPAPIEKTISKAK 360
    |||

Qy  122 GQPREPQVYTLPPSRDELTKNQVSLTCLVKGFYPSDIAVEWESNGQPENNYKTTTPPVLD 181
    |||
Db  361 GQPREPQVYTLPPSRDELTKNQVSLTCLVKGFYPSDIAVEWESNGQPENNYKTTTPPVLD 420
    |||

Qy  182 DGSFFLYSKLTVDKSRWQOGNVFSCSVMHAEALHNHYTQKSLSLSPGKDWLK--AF----- 234
    |||
Db  421 DGSFFLYSKLTVDKSRWQOGNVFSCSVMHAEALHNHYTQKSLSLSPGSGGLKIAAFNIQT 480
    |||

Qy  235 -----YDKV----- 234
Db  481 GETKMSNATLVSYIVQILSRVDIALVQEVDRSHLTAVGKLLDNLNQADPOTHYVYVSEPL 540
    |||

Qy  235 -----YDKV----- 234
Db  541 GRNSYKERYLFVYRPDQVSAVDSYYDDGCEPCGNDTFNREPAIVRFSETEVREFAIV 600
    |||

Qy  235 -----YDKV----- 234
Db  601 PLHAAPGDAVAEIDALYDVLDVQEKWGLDVLMGLDFNAGCSYVRPSQWSSIRLWTSPT 660
    |||

Qy  235 -----YDKV----- 234
Db  661 FQWLIPDSADTTPPTHCAYDRIVVAGMLLRGAVVPDSALFFNFQAAIYGLSDQLAAQ 717
    |||

RESULT 5
US-09-840-277-2
; Sequence 2, Application US/09840277
; Patent No. US20020168363A1
; GENERAL INFORMATION:
; APPLICANT: FEIGE, ULRICH
; APPLICANT: KOHNO, TADAHIKO
; APPLICANT: LACEY, DAVID LEE
; APPLICANT: BOONE, THOMAS CHARLES
; TITLE OF INVENTION: INTEGRIN/ADHESION ANTAGONISTS
; FILE REFERENCE: A-688A
; CURRENT APPLICATION NUMBER: US/09/840,277
; CURRENT FILING DATE: 2001-08-14
; PRIOR APPLICATION NUMBER: 60/198,919
; PRIOR FILING DATE: 2000-04-21
; PRIOR APPLICATION NUMBER: 60/201,394
; PRIOR FILING DATE: 2000-05-03
```

```
; NUMBER OF SEQ ID NOS: 135
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 2
; LENGTH: 228
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-840-277-2

Query Match          92.8%; Score 1238; DB 9; Length 228;
Best Local Similarity 100.0%; Pred. No. 4e-33;
Matches 228; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy  1 MDKTHCTPCPAPPELLGGPSVFLFPPKPKDTLMISRTPEVTCVVVDVSHEDPEVKFNMYV 60
    |||
Db  1 MDKTHCTPCPAPPELLGGPSVFLFPPKPKDTLMISRTPEVTCVVVDVSHEDPEVKFNMYV 60
    |||

Qy  61 DGVEVHNAKTKPREEQYNSTYRVVSVLTVLHQDWLNGKEYKCKVSNKALPAPIEKTISKA 120
    |||
Db  61 DGVEVHNAKTKPREEQYNSTYRVVSVLTVLHQDWLNGKEYKCKVSNKALPAPIEKTISKA 120
    |||

Qy  121 KGQPREPQVYTLPPSRDELTKNQVSLTCLVKGFYPSDIAVEWESNGQPENNYKTTTPVLD 180
    |||
Db  121 KGQPREPQVYTLPPSRDELTKNQVSLTCLVKGFYPSDIAVEWESNGQPENNYKTTTPVLD 180
    |||

Qy  181 SDGSFFLYSKLTVDKSRWQOGNVFSCSVMHAEALHNHYTQKSLSLSPGK 228
    |||
Db  181 SDGSFFLYSKLTVDKSRWQOGNVFSCSVMHAEALHNHYTQKSLSLSPGK 228
    |||

RESULT 6
US-09-847-249A-2
; Sequence 2, Application US/09847249A
; Publication No. US20030032588A1
; GENERAL INFORMATION:
; APPLICANT: MARSHALL, WILLIAM S.
; APPLICANT: STARK, KEVIN LEE
; TITLE OF INVENTION: GLUCAGON ANTAGONIST
; FILE REFERENCE: A-693
; CURRENT APPLICATION NUMBER: US/09/847,249A
; CURRENT FILING DATE: 2001-05-02
; PRIOR APPLICATION NUMBER: 60/201,436
; PRIOR FILING DATE: 2000-05-03
; NUMBER OF SEQ ID NOS: 80
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 2
; LENGTH: 228
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-847-249A-2

Query Match          92.8%; Score 1238; DB 9; Length 228;
Best Local Similarity 100.0%; Pred. No. 4e-33;
Matches 228; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy  1 MDKTHCTPCPAPPELLGGPSVFLFPPKPKDTLMISRTPEVTCVVVDVSHEDPEVKFNMYV 60
    |||
Db  1 MDKTHCTPCPAPPELLGGPSVFLFPPKPKDTLMISRTPEVTCVVVDVSHEDPEVKFNMYV 60
    |||

Qy  61 DGVEVHNAKTKPREEQYNSTYRVVSVLTVLHQDWLNGKEYKCKVSNKALPAPIEKTISKA 120
    |||
Db  61 DGVEVHNAKTKPREEQYNSTYRVVSVLTVLHQDWLNGKEYKCKVSNKALPAPIEKTISKA 120
    |||

Qy  121 KGQPREPQVYTLPPSRDELTKNQVSLTCLVKGFYPSDIAVEWESNGQPENNYKTTTPVLD 180
    |||
Db  121 KGQPREPQVYTLPPSRDELTKNQVSLTCLVKGFYPSDIAVEWESNGQPENNYKTTTPVLD 180
    |||

Qy  181 SDGSFFLYSKLTVDKSRWQOGNVFSCSVMHAEALHNHYTQKSLSLSPGK 228
    |||
Db  181 SDGSFFLYSKLTVDKSRWQOGNVFSCSVMHAEALHNHYTQKSLSLSPGK 228
    |||

RESULT 7
US-09-840-669B-2
```

Db 421 DGSFFLYSKLTVDKSRWQGNVFCVSMHEALHNHYTQKSLSLSPCKGSGGLKIAAFNIQ 480  
QY 229 -----KAFYDKV-----AEKLKEA 245  
Db 481 TFGETKMSNATLVSYIVQILSRVDIALVOEVRDLSHTAVGKLLDNLNQADPTDTHYVYVSE 540  
QY 229 -----KAFYDKV-----AEKLKEA 245  
Db 541 PLGRNSYKERYLFVYRPDQVSAVDSYDDGCEPCGNDTFNREPAIVRFFSRFTEVREFA 600  
QY 229 -----KAFYDKV-----AEKLKEA 245  
Db 601 IVP LHAAPGDAVAEIDALYDVYLDVQEKWGLDVMLMGDFNAGCCSYVRPSQSSIRLWTS 660  
QY 229 -----KAFYDKV-----AEKLKEA 245  
Db 661 PTFQWLIPDSADTTATPTHCAYDRIVVAGMLLRGAVVPSALPFFNQAAAYGLSDQLAQ 719  
QY 229 -----KAFYDKV-----AEKLKEA 245  
RESULT 2  
US-09-825-012-55  
; Sequence 55, Application US/09825012  
; Patent No. US20020122798A1  
; GENERAL INFORMATION:  
; APPLICANT: Young, Robert  
; TITLE OF INVENTION: Compounds for Targeting  
; FILE REFERENCE: 43191-256808  
; CURRENT APPLICATION NUMBER: US/09/825,012  
; CURRENT FILING DATE: 2001-04-03  
; PRIOR APPLICATION NUMBER: US 60/237,159  
; PRIOR FILING DATE: 2000-10-02  
; PRIOR APPLICATION NUMBER: GB 0008049.9  
; NUMBER OF SEQ ID NOS: 102  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 55  
; LENGTH: 741  
; TYPE: PRT  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Humanised HMFGL heavy chain - DNase I fusion  
US-09-825-012-55  
Query Match 93.4%; Score 1246; DB 10; Length 741;  
Best Local Similarity 48.6%; Pred. No. 3.4e-32;  
Matches 233; Conservative 6; Mismatches 5; Indels 235; Gaps 3;  
QY 2 DKTHTCCPCPAPPELLGGPSVFLFPPKPKDTLMISRTPEVTCVVDVSHEDPEVKFNWYVD 61  
Db 241 DKTHTCCPCPAPPELLGGPSVFLFPPKPKDTLMISRTPEVTCVVDVSHEDPEVKFNWYVD 300  
QY 62 GVEVHNAKTPREEQNSTYRVSVLTVLHODWLNCKEYKCKVSNKALPAPIETKISKAK 121  
Db 301 GVEVHNAKTPREEQNSTYRVSVLTVLHODWLNCKEYKCKVSNKALPAPIETKISKAK 360  
QY 122 GPREPQVYTLPPSRDELTKNQVSLTCLVKGFYPSDIAVEWESNGOPENNYKTTTPVLD 181  
Db 361 GPREPQVYTLPPSRDELTKNQVSLTCLVKGFYPSDIAVEWESNGOPENNYKTTTPVLD 420  
QY 182 DGSFFLYSKLTVDKSRWQGNVFCVSMHEALHNHYTQKSLSLSPCKGSGGLKIAAFNIQ 480  
Db 421 DGSFFLYSKLTVDKSRWQGNVFCVSMHEALHNHYTQKSLSLSPCKGSGGLKIAAFNIQ 540  
QY 229 -----KAFYDKV-----AEKLKEA 245  
Db 481 TFGETKMSNATLVSYIVQILSRVDIALVOEVRDLSHTAVGKLLDNLNQADPTDTHYVYVSE 540  
QY 229 -----KAFYDKV-----AEKLKEA 245  
Db 541 PLGRNSYKERYLFVYRPDQVSAVDSYDDGCEPCGNDTFNREPAIVRFFSRFTEVREFA 600  
QY 229 -----KAFYDKV-----AEKLKEA 245

Db 601 IVP LHAAPGDAVAEIDALYDVYLDVQEKWGLDVMLMGDFNAGCCSYVRPSQSSIRLWTS 660  
QY 229 -----KAFYDKV-----AEKLKEA 245  
Db 661 PTFQWLIPDSADTTATPTHCAYDRIVVAGMLLRGAVVPSALPFFNQAAAYGLSDQLAQ 719  
QY 229 -----KAFYDKV-----AEKLKEA 245  
RESULT 3  
US-09-825-012-52  
; Sequence 52, Application US/09825012  
; Patent No. US20020122798A1  
; GENERAL INFORMATION:  
; APPLICANT: Young, Robert  
; TITLE OF INVENTION: Compounds for Targeting  
; FILE REFERENCE: 43191-256808  
; CURRENT APPLICATION NUMBER: US/09/825,012  
; CURRENT FILING DATE: 2001-04-03  
; PRIOR APPLICATION NUMBER: US 60/237,159  
; PRIOR FILING DATE: 2000-10-02  
; PRIOR APPLICATION NUMBER: GB 0008049.9  
; NUMBER OF SEQ ID NOS: 102  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 52  
; LENGTH: 729  
; TYPE: PRT  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Humanised HMFGL heavy chain - DNase I fusion  
US-09-825-012-52  
Query Match 93.3%; Score 1245; DB 10; Length 729;  
Best Local Similarity 49.1%; Pred. No. 3.5e-32;  
Matches 234; Conservative 6; Mismatches 4; Indels 233; Gaps 3;  
QY 2 DKTHTCCPCPAPPELLGGPSVFLFPPKPKDTLMISRTPEVTCVVDVSHEDPEVKFNWYVD 61  
Db 241 DKTHTCCPCPAPPELLGGPSVFLFPPKPKDTLMISRTPEVTCVVDVSHEDPEVKFNWYVD 300  
QY 62 GVEVHNAKTPREEQNSTYRVSVLTVLHODWLNCKEYKCKVSNKALPAPIETKISKAK 121  
Db 301 GVEVHNAKTPREEQNSTYRVSVLTVLHODWLNCKEYKCKVSNKALPAPIETKISKAK 360  
QY 122 GPREPQVYTLPPSRDELTKNQVSLTCLVKGFYPSDIAVEWESNGOPENNYKTTTPVLD 181  
Db 361 GPREPQVYTLPPSRDELTKNQVSLTCLVKGFYPSDIAVEWESNGOPENNYKTTTPVLD 420  
QY 182 DGSFFLYSKLTVDKSRWQGNVFCVSMHEALHNHYTQKSLSLSPCKGSGGLKIAAFNIQ 480  
Db 421 DGSFFLYSKLTVDKSRWQGNVFCVSMHEALHNHYTQKSLSLSPCKGSGGLKIAAFNIQ 540  
QY 235 -----YDKV-----AEKLKEA 245  
Db 481 GETKMSNATLVSYIVQILSRVDIALVOEVRDLSHTAVGKLLDNLNQADPTDTHYVYVSE 540  
QY 235 -----YDKV-----AEKLKEA 245  
Db 541 GRNSYKERYLFVYRPDQVSAVDSYDDGCEPCGNDTFNREPAIVRFFSRFTEVREFA 600  
QY 235 -----YDKV-----AEKLKEA 245  
Db 601 PLHAAPGDAVAEIDALYDVYLDVQEKWGLDVMLMGDFNAGCCSYVRPSQSSIRLWTSPT 660  
QY 235 -----YDKV-----AEKLKEA 245  
Db 661 FQWLIPDSADTTATPTHCAYDRIVVAGMLLRGAVVPSALPFFNQAAAYGLSDQLAQ 717  
QY 235 -----YDKV-----AEKLKEA 245  
RESULT 4  
US-09-825-012-61  
; Sequence 61, Application US/09825012  
; Patent No. US20020122798A1  
; GENERAL INFORMATION:

GenCore version 5.1.4.p5.4578  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: April 21, 2003, 10:43:21 ; Search time 11.3353 seconds  
(without alignments)  
1640.982 Million cell updates/sec

Title: 2LINK7  
Perfect score: 1334  
Sequence: 1 MDKTHCTPCPAPELLGGPS.....GKDLKAFYDKVAEXLKEAF 246

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.0

Searched: 288829 seqs, 75613885 residues  
Total number of hits satisfying chosen parameters: 288829

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

- Database : Published\_Applications\_AA:\*
- 1: /cgn2\_6/ptodata/2/pubpaa/US08\_NEW\_PUB.pep.\*
  - 2: /cgn2\_6/ptodata/2/pubpaa/PCT\_NEW\_PUB.pep.\*
  - 3: /cgn2\_6/ptodata/2/pubpaa/US06\_NEW\_PUB.pep.\*
  - 4: /cgn2\_6/ptodata/2/pubpaa/US06\_PUBCOMB.pep.\*
  - 5: /cgn2\_6/ptodata/2/pubpaa/US07\_NEW\_PUB.pep.\*
  - 6: /cgn2\_6/ptodata/2/pubpaa/US07\_PUBCOMB.pep.\*
  - 7: /cgn2\_6/ptodata/2/pubpaa/PCTUS\_PUBCOMB.pep.\*
  - 8: /cgn2\_6/ptodata/2/pubpaa/US08\_PUBCOMB.pep.\*
  - 9: /cgn2\_6/ptodata/2/pubpaa/US09\_NEW\_PUB.pep.\*
  - 10: /cgn2\_6/ptodata/2/pubpaa/US03\_PUBCOMB.pep.\*
  - 11: /cgn2\_6/ptodata/2/pubpaa/US10\_NEW\_PUB.pep.\*
  - 12: /cgn2\_6/ptodata/2/pubpaa/US10\_PUBCOMB.pep.\*
  - 13: /cgn2\_6/ptodata/2/pubpaa/US60\_NEW\_PUB.pep.\*
  - 14: /cgn2\_6/ptodata/2/pubpaa/US60\_PUBCOMB.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1246	93.4	731	10	US-09-825-012-46
2	1246	93.4	741	10	US-09-825-012-55
3	1245	93.3	729	10	US-09-825-012-52
4	1245	93.3	739	10	US-09-825-012-61
5	1238	92.8	228	9	US-09-840-277-2
6	1238	92.8	228	9	US-09-847-249A-2
7	1238	92.8	228	9	US-09-840-669B-2
8	1238	92.8	228	9	US-09-843-221A-2
9	1238	92.8	228	10	US-09-847-712-2
10	1238	92.8	282	9	US-09-840-277-109
11	1238	92.8	389	10	US-09-784-623-14
12	1237	92.7	394	10	US-09-854-864-31
13	1237	92.7	730	10	US-09-825-012-49
14	1237	92.7	740	10	US-09-825-012-58
15	1234	92.5	229	9	US-10-215-297-2
16	1234	92.5	229	9	US-10-215-298-2
17	1234	92.5	347	9	US-10-091-236-17
18	1234	92.5	347	9	US-10-091-313-7
19	1234	92.5	347	9	US-10-091-268-7

ALIGNMENTS

RESULT 1									
US-09-825-012-46									
; Sequence 46, Application US/09825012									
; Patent No. US20020122798A1									
; GENERAL INFORMATION:									
; APPLICANT: Young, Robert									
; TITLE OF INVENTION: Compounds for Targeting									
; FILE REFERENCE: 43191-256808									
; CURRENT APPLICATION NUMBER: US/09/825,012									
; CURRENT FILING DATE: 2001-04-03									
; PRIOR APPLICATION NUMBER: US 60/237,159									
; PRIOR FILING DATE: 2000-10-02									
; PRIOR APPLICATION NUMBER: GB 0008049.9									
; PRIOR FILING DATE: 2000-04-03									
; NUMBER OF SEQ ID NOS: 102									
; SOFTWARE: Patentin version 3.1									
; SEQ ID NO 46									
; LENGTH: 731									
; TYPE: PRT									
; ORGANISM: Artificial Sequence									
; FEATURE:									
; OTHER INFORMATION: Humanised HMFGL heavy chain - DNase I fusion									
US-09-825-012-46									
Query Match 93.4%; Score 1246; DB 10; Length 731;									
Best Local Similarity 48.6%; Pred. No. 3 3e-32;									
Matches 233; Conservative 6; Mismatches 5; Indels 235; Gaps 3;									
QY 2 DKTHCTPCPAPELLGGPSVFLFPKPKDPLMISRPEVTCVVVDVSHEDPEVKFNMYVD 61									
DB 241 DKTHCTPCPAPELLGGPSVFLFPKPKDPLMISRPEVTCVVVDVSHEDPEVKFNMYVD 300									
QY 62 GVEVHNAKTPREQNSTYRVVSVLTVLHODWLNKKEYCKYKSNKALPAPIETISKAK 121									
DB 301 GVEVHNAKTPREQNSTYRVVSVLTVLHODWLNKKEYCKYKSNKALPAPIETISKAK 360									
QY 122 GQREPVVTLPSRDELTKNQVSLTCLVKGFYPSDIAVEWESNGQPENNYKTTTPVLD 181									
DB 361 GQREPVVTLPSRDELTKNQVSLTCLVKGFYPSDIAVEWESNGQPENNYKTTTPVLD 420									
QY 182 DGSFFLYSLKTVTDKSRWQGNFSCVMHEALHNHTVTKSLSPGK----- 228									

Sequence 8, Appl  
Sequence 8, Appl  
Sequence 36, Appl  
Sequence 39, Appl  
Sequence 8, Appl  
Sequence 9, Appl  
Sequence 2, Appl  
Sequence 42, Appl  
Sequence 44, Appl  
Sequence 2, Appl  
Sequence 22, Appl  
Sequence 10, Appl  
Sequence 6, Appl  
Sequence 13, Appl  
Sequence 12, Appl  
Sequence 10, Appl  
Sequence 9, Appl  
Sequence 15, Appl  
Sequence 11, Appl  
Sequence 22, Appl  
Sequence 16, Appl  
Sequence 18, Appl  
Sequence 16, Appl  
Sequence 5, Appl  
Sequence 17, Appl



Matches 227;		Conservative	0;	Mismatches	0;	Indels	0;	Gaps	0;
Qy	2	DKTHTCPCPAPPELLGGPSVFLFPKPKDTLMIISRTPEVTCVVVDVSHEDPEVKENWYVD							
Db	150	DKTHTCPCPAPPELLGGPSVFLFPKPKDTLMIISRTPEVTCVVVDVSHEDPEVKENWYVD							
Qy	62	GVEVHNAKTKPREQYNSTYRVVSVLTVLHQDWLNGKEYCKVSNKALPAPIEKTISKAK							
Db	210	GVEVHNAKTKPREQYNSTYRVVSVLTVLHQDWLNGKEYCKVSNKALPAPIEKTISKAK							
Qy	122	GOPREPOVYTLPPSRDELTKNQVSLTCLVKGFYPSDIAVEWESNGQPENNYKTTTPPVLD							
Db	270	GOPREPOVYTLPPSRDELTKNQVSLTCLVKGFYPSDIAVEWESNGQPENNYKTTTPPVLD							
Qy	182	DGSFFLYSKLTVDKSRWQOGNVFSCVMHEALHNHYTOKLSLSLSPGK							
Db	330	DGSFFLYSKLTVDKSRWQOGNVFSCVMHEALHNHYTOKLSLSLSPGK							

Search completed: April 21, 2003, 10:50:21  
Job time : 14.7471 secs

APPLICATION NUMBER: 07/936190  
FILING DATE: 26-AUG-1992  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 07/842777  
FILING DATE: 18-FEB-1992  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 07/250785  
FILING DATE: 28-SEP-1988  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 07/104329  
FILING DATE: 02-OCT-1987  
ATTORNEY/AGENT INFORMATION:  
NAME: Hasak, Janet E.  
REGISTRATION NUMBER: 28,616  
REFERENCE/DOCKET NUMBER: 444P1C2  
TELEPHONE: 415/225-1896  
TELEFAX: 415/952-9881  
TELEX: 910/371-7168  
INFORMATION FOR SEQ ID NO: 7:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 371 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
US-08-236-311-7

Query Match 92.4%; Score 1233; DB 1; Length 371;  
Best Local Similarity 100.0%; Pred. No. 1.4e-42;  
Matches 227; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 2 DKHTCPCPAPPELLGGPSVFLPPKPKDTLMISRTPEVTCVVDVSHEDPEVKFNWYVD 61  
DB 145 DKHTCPCPAPPELLGGPSVFLPPKPKDTLMISRTPEVTCVVDVSHEDPEVKFNWYVD 204  
QY 62 GVEVHNAKTPREEQNSTYRVVSVLTVLHQDWLNGKEYCKVSNKALPAPIEKTISKAK 121  
DB 205 GVEVHNAKTPREEQNSTYRVVSVLTVLHQDWLNGKEYCKVSNKALPAPIEKTISKAK 264  
QY 122 GQPREPQVYTLPPSRDELTKNQVSLTCLVKGFYPSDIAVEWESNGOPENNYKTTTPVLD 181  
DB 265 GQPREPQVYTLPPSRDELTKNQVSLTCLVKGFYPSDIAVEWESNGOPENNYKTTTPVLD 324  
QY 182 DGSFFLYSKLTVDKSRWQOGNFSVCSVMHEALHNHYTQKSLSLSPGK 228  
DB 325 DGSFFLYSKLTVDKSRWQOGNFSVCSVMHEALHNHYTQKSLSLSPGK 371

RESULT 14  
US-08-457-918-7  
Sequence 7, Application US/08457918  
Patent No. 6117655  
GENERAL INFORMATION:  
APPLICANT: Capon, Daniel J.  
APPLICANT: Gregory, Timothy J.  
TITLE OF INVENTION: Adhesion Variants  
NUMBER OF SEQUENCES: 25  
CORRESPONDENCE ADDRESS:  
ADDRESSER: Genentech, Inc.  
STREET: 460 Point San Bruno Blvd  
CITY: South San Francisco  
STATE: California  
COUNTRY: USA  
ZIP: 94080  
COMPUTER READABLE FORM:  
MEDIUM TYPE: 5.25 inch, 360 Kb floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: patin (Genentech)  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/457,918  
FILING DATE: 1-JUN-1995  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:

APPLICATION NUMBER: 08/236311  
FILING DATE: 02-MAY-1994  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 07/936190  
FILING DATE: 26-AUG-1992  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 07/842777  
FILING DATE: 18-FEB-1992  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 07/250785  
FILING DATE: 28-SEP-1988  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 07/104329  
FILING DATE: 02-OCT-1987  
ATTORNEY/AGENT INFORMATION:  
NAME: Kudinec, Jeffrey S.  
REGISTRATION NUMBER: 36,575  
REFERENCE/DOCKET NUMBER: P0444P1C3  
TELEPHONE: 415/225-8228  
TELEFAX: 415/952-9881  
TELEX: 910/371-7168  
INFORMATION FOR SEQ ID NO: 7:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 371 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
US-08-457-918-7

Query Match 92.4%; Score 1233; DB 3; Length 371;  
Best Local Similarity 100.0%; Pred. No. 1.4e-42;  
Matches 227; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 2 DKHTCPCPAPPELLGGPSVFLPPKPKDTLMISRTPEVTCVVDVSHEDPEVKFNWYVD 61  
DB 145 DKHTCPCPAPPELLGGPSVFLPPKPKDTLMISRTPEVTCVVDVSHEDPEVKFNWYVD 204  
QY 62 GVEVHNAKTPREEQNSTYRVVSVLTVLHQDWLNGKEYCKVSNKALPAPIEKTISKAK 121  
DB 205 GVEVHNAKTPREEQNSTYRVVSVLTVLHQDWLNGKEYCKVSNKALPAPIEKTISKAK 264  
QY 122 GQPREPQVYTLPPSRDELTKNQVSLTCLVKGFYPSDIAVEWESNGOPENNYKTTTPVLD 181  
DB 265 GQPREPQVYTLPPSRDELTKNQVSLTCLVKGFYPSDIAVEWESNGOPENNYKTTTPVLD 324  
QY 182 DGSFFLYSKLTVDKSRWQOGNFSVCSVMHEALHNHYTQKSLSLSPGK 228  
DB 325 DGSFFLYSKLTVDKSRWQOGNFSVCSVMHEALHNHYTQKSLSLSPGK 371

RESULT 15  
US-09-180-100-22  
Sequence 22, Application US/09180100  
Patent No. 6306395  
GENERAL INFORMATION:  
APPLICANT: NAKAMURA, No. 630639510  
APPLICANT: NAGATA, Shigekazu  
TITLE OF INVENTION: NOVEL Fas ANTIGEN DERIVATIVE  
FILE REFERENCE: 1110-207P  
CURRENT APPLICATION NUMBER: US/09/180,100  
CURRENT FILING DATE: 1998-11-02  
EARLIER APPLICATION NUMBER: PCT/JP97/01502  
EARLIER FILING DATE: 1997-05-01  
NUMBER OF SEQ ID NOS: 25  
SOFTWARE: Patent In Ver. 2.0  
SEQ ID NO 22  
LENGTH: 376  
TYPE: PRT  
ORGANISM: Homo sapiens  
US-09-180-100-22  
Query Match 92.4%; Score 1233; DB 4; Length 376;  
Best Local Similarity 100.0%; Pred. No. 1.4e-42;

;; CURRENT APPLICATION NUMBER: US/09/131,247  
;; CURRENT FILING DATE: 1998-08-07  
;; EARLIER APPLICATION NUMBER: 60/055,185  
;; EARLIER FILING DATE: 1997-08-08  
;; EARLIER APPLICATION NUMBER: PCT/US 97/02131  
;; EARLIER FILING DATE: 1997-02-10  
;; NUMBER OF SEQ ID NOS: 16  
;; SOFTWARE: PatentIn Ver. 2.0  
;; SEQ ID NO 6  
;; LENGTH: 235  
;; TYPE: PRT  
;; ORGANISM: Human  
US-09-131-247-6

Query Match 92.4%; Score 1233; DB 4; Length 235;  
Best Local Similarity 100.0%; Pred. No. 4.3e-43;  
Matches 227; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 2 DKHTCCPCPAPPELLGGPSVFLFPPKPKDTLMISRTPEVTCVVDVSHEDPEVKENWYVD 61  
DB 9 DKHTCCPCPAPPELLGGPSVFLFPPKPKDTLMISRTPEVTCVVDVSHEDPEVKENWYVD 68  
QY 62 GVEVHNAKTKPREEQYNSTYRVVSVLTVLHQDWLNGKEYCKVSNKALPAPIEKTISKAK 121  
DB 69 GVEVHNAKTKPREEQYNSTYRVVSVLTVLHQDWLNGKEYCKVSNKALPAPIEKTISKAK 128  
QY 122 GQPREPQVYTLPPSRDELTKNQVSLTCLVKGFYPSDIAVWESNGQPENNYKTTPPVLD 181  
DB 129 GQPREPQVYTLPPSRDELTKNQVSLTCLVKGFYPSDIAVWESNGQPENNYKTTPPVLD 188  
QY 182 DGSFELYSLKTVDKSRWQGNVFCSCVMHEALHNYHTOKSLSPGK 228  
DB 189 DGSFELYSLKTVDKSRWQGNVFCSCVMHEALHNYHTOKSLSPGK 235

RESULT 11  
US-09-178-869-2  
;; Sequence 2, Application US/09178869B  
;; Patent No. 6197294  
;; GENERAL INFORMATION:  
;; APPLICANT: Tao, Weng  
;; APPLICANT: Wong, Shou  
;; APPLICANT: Hickey, William F.  
;; APPLICANT: Hamming, Joseph P.  
;; APPLICANT: Baetge, E. Edward  
;; TITLE OF INVENTION: CELL SURFACE-INDUCED MACROPHAGE ACTIVATION  
;; FILE REFERENCE: 17810-043  
;; CURRENT APPLICATION NUMBER: US/09/178,869B  
;; CURRENT FILING DATE: 1998-10-26  
;; NUMBER OF SEQ ID NOS: 14  
;; SOFTWARE: PatentIn Ver. 2.0  
;; SEQ ID NO 2  
;; LENGTH: 331  
;; TYPE: PRT  
;; ORGANISM: Homo sapiens  
US-09-178-869-2

Query Match 92.4%; Score 1233; DB 4; Length 331;  
Best Local Similarity 100.0%; Pred. No. 1e-42;  
Matches 227; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 2 DKHTCCPCPAPPELLGGPSVFLFPPKPKDTLMISRTPEVTCVVDVSHEDPEVKENWYVD 61  
DB 105 DKHTCCPCPAPPELLGGPSVFLFPPKPKDTLMISRTPEVTCVVDVSHEDPEVKENWYVD 164  
QY 62 GVEVHNAKTKPREEQYNSTYRVVSVLTVLHQDWLNGKEYCKVSNKALPAPIEKTISKAK 121  
DB 165 GVEVHNAKTKPREEQYNSTYRVVSVLTVLHQDWLNGKEYCKVSNKALPAPIEKTISKAK 224  
QY 122 GQPREPQVYTLPPSRDELTKNQVSLTCLVKGFYPSDIAVWESNGQPENNYKTTPPVLD 181  
DB 225 GQPREPQVYTLPPSRDELTKNQVSLTCLVKGFYPSDIAVWESNGQPENNYKTTPPVLD 284

QY 182 DGSFELYSLKTVDKSRWQGNVFCSCVMHEALHNYHTOKSLSPGK 228  
DB 285 DGSFELYSLKTVDKSRWQGNVFCSCVMHEALHNYHTOKSLSPGK 331  
RESULT 12  
US-09-180-100-11  
;; Sequence 11, Application US/09180100  
;; Patent No. 6306395  
;; GENERAL INFORMATION:  
;; APPLICANT: NAKAMURA, No. 630639510  
;; APPLICANT: NAGATA, Shigekazu  
;; TITLE OF INVENTION: NOVEL Fas ANTIGEN DERIVATIVE  
;; FILE REFERENCE: 1110-207P  
;; CURRENT APPLICATION NUMBER: US/09/180,100  
;; CURRENT FILING DATE: 1998-11-02  
;; EARLIER APPLICATION NUMBER: PCT/JP97/01502  
;; EARLIER FILING DATE: 1997-05-01  
;; NUMBER OF SEQ ID NOS: 25  
;; SOFTWARE: PatentIn Ver. 2.0  
;; SEQ ID NO 11  
;; LENGTH: 360  
;; TYPE: PRT  
;; ORGANISM: Homo sapiens  
US-09-180-100-11

Query Match 92.4%; Score 1233; DB 4; Length 360;  
Best Local Similarity 100.0%; Pred. No. 1.3e-42;  
Matches 227; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 2 DKHTCCPCPAPPELLGGPSVFLFPPKPKDTLMISRTPEVTCVVDVSHEDPEVKENWYVD 61  
DB 134 DKHTCCPCPAPPELLGGPSVFLFPPKPKDTLMISRTPEVTCVVDVSHEDPEVKENWYVD 193  
QY 62 GVEVHNAKTKPREEQYNSTYRVVSVLTVLHQDWLNGKEYCKVSNKALPAPIEKTISKAK 121  
DB 194 GVEVHNAKTKPREEQYNSTYRVVSVLTVLHQDWLNGKEYCKVSNKALPAPIEKTISKAK 253  
QY 122 GQPREPQVYTLPPSRDELTKNQVSLTCLVKGFYPSDIAVWESNGQPENNYKTTPPVLD 181  
DB 254 GQPREPQVYTLPPSRDELTKNQVSLTCLVKGFYPSDIAVWESNGQPENNYKTTPPVLD 313  
QY 182 DGSFELYSLKTVDKSRWQGNVFCSCVMHEALHNYHTOKSLSPGK 228  
DB 314 DGSFELYSLKTVDKSRWQGNVFCSCVMHEALHNYHTOKSLSPGK 360

RESULT 13  
US-08-236-311-7  
;; Sequence 7, Application US/08236311  
;; Patent No. 5565335  
;; GENERAL INFORMATION:  
;; APPLICANT: Capon, Daniel J.  
;; APPLICANT: Gregory, Timothy J.  
;; TITLE OF INVENTION: Adhesion Variants  
;; NUMBER OF SEQUENCES: 25  
;; CORRESPONDENCE ADDRESS:  
;; ADDRESSEE: Genentech, Inc.  
;; STREET: 460 Point San Bruno Blvd  
;; CITY: South San Francisco  
;; STATE: California  
;; COUNTRY: USA  
;; ZIP: 94080  
COMPUTER READABLE FORM:  
MEDIUM TYPE: 5.25 inch, 360 Kb floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: patin (genentech)  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/236,311  
FILING DATE: 02-MAY-1994  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:

Db 300 SDGSFFLYSKLTVDKSRWQGNVFCVSMHEALHNNHYTKSLSPGK 347

RESULT 8  
US-09-189-129-2  
; Sequence 2, Application US/09189129  
; Patent No. 6323027  
; GENERAL INFORMATION:  
; APPLICANT: Burkly, Linda C  
; APPLICANT: Benjamin, Christopher D  
; APPLICANT: Hession, Catherine A  
; APPLICANT: Whitty, Adrian  
; TITLE OF INVENTION: COMMON GAMMA CHAIN BLOCKING AGENTS  
; NUMBER OF SEQUENCES: 17  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Biogen, Inc.  
; STREET: 14 Cambridge Center  
; CITY: Cambridge  
; STATE: Massachusetts  
; COUNTRY: USA  
; ZIP: 02142

COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/09/189,129  
; FILING DATE:  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: A006 PCT CIP  
; FILING DATE: 09-MAY-1997  
; APPLICATION NUMBER: 60/017,466  
; FILING DATE: 10-MAY-1996  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Kaplan, Warren A.  
; REGISTRATION NUMBER: 34,199  
; REFERENCE/DOCKET NUMBER: A006 PCT CIP  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 617 679-2838  
; TELEFAX: 617 679-2838  
; INFORMATION FOR SEQ ID NO: 2:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 482 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: YES  
; HYPOTHETICAL: YES  
; ANTI-SENSE: NO  
US-09-189-129-2

Query Match 92.5%; Score 1234; DB 4; Length 482;  
Best Local Similarity 99.6%; Pred. No. 2.5e-42;  
Matches 227; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 MKTHTCCPCPAPPELLGGPSVFLFPPKPKDTLMISRTPEVTCVVDVSHEDPEVKFNWYV 60  
Db 255 VDKTHTCCPCPAPPELLGGPSVFLFPPKPKDTLMISRTPEVTCVVDVSHEDPEVKFNWYV 314  
QY 61 DGEVHNATKPREQYNSYRVSVLTVLHQDLNGLNGKEYCKVSNKALPAPIETKISK 120  
Db 315 DGEVHNATKPREQYNSYRVSVLTVLHQDLNGLNGKEYCKVSNKALPAPIETKISK 374  
QY 121 KGOPREPVYTLPPSRDELTKNQVSLTCLVKGFYPSDIAVEWESNGQPENNYKTTPPVLD 180  
Db 375 KGOPREPVYTLPPSRDELTKNQVSLTCLVKGFYPSDIAVEWESNGQPENNYKTTPPVLD 434  
QY 181 SDGSFFLYSKLTVDKSRWQGNVFCVSMHEALHNNHYTKSLSPGK 228  
Db 435 SDGSFFLYSKLTVDKSRWQGNVFCVSMHEALHNNHYTKSLSPGK 482

RESULT 9  
US-08-595-043A-50  
; Sequence 50, Application US/08595043A  
; Patent No. 5935824  
; GENERAL INFORMATION:  
; APPLICANT: SCARLATO, GREGORY D.  
; TITLE OF INVENTION: PROTEIN EXPRESSION SYSTEM  
; NUMBER OF SEQUENCES: 90  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: MEDLEN & CARROLL  
; STREET: 220 MONTGOMERY STREET, SUITE 2200  
; CITY: SAN FRANCISCO  
; STATE: CALIFORNIA  
; COUNTRY: UNITED STATES OF AMERICA  
; ZIP: 94104

COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/595,043A  
; FILING DATE: 31-JAN-1996  
; CLASSIFICATION: 435  
; ATTORNEY/AGENT INFORMATION:  
; NAME: CARROLL, PETER G.  
; REGISTRATION NUMBER: 32,837  
; REFERENCE/DOCKET NUMBER: SGAR-00371  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (415) 705-8410  
; TELEFAX: (415) 397-8338  
; INFORMATION FOR SEQ ID NO: 50:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 232 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
US-08-595-043A-50

Query Match 92.4%; Score 1233; DB 2; Length 232;  
Best Local Similarity 100.0%; Pred. No. 4.2e-43;  
Matches 227; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 DKTHTCPCPAPPELLGGPSVFLFPPKPKDTLMISRTPEVTCVVDVSHEDPEVKFNWYVD 61  
Db 6 DKTHTCPCPAPPELLGGPSVFLFPPKPKDTLMISRTPEVTCVVDVSHEDPEVKFNWYVD 65  
QY 62 GVEVHNATKPREQYNSYRVSVLTVLHQDLNGLNGKEYCKVSNKALPAPIETKISK 121  
Db 66 GVEVHNATKPREQYNSYRVSVLTVLHQDLNGLNGKEYCKVSNKALPAPIETKISK 125  
QY 122 GGPPEPVYTLPPSRDELTKNQVSLTCLVKGFYPSDIAVEWESNGQPENNYKTTPPVLD 181  
Db 126 GGPPEPVYTLPPSRDELTKNQVSLTCLVKGFYPSDIAVEWESNGQPENNYKTTPPVLD 185  
QY 182 DGSFFLYSKLTVDKSRWQGNVFCVSMHEALHNNHYTKSLSPGK 228  
Db 186 DGSFFLYSKLTVDKSRWQGNVFCVSMHEALHNNHYTKSLSPGK 232

RESULT 10  
US-09-131-247-6  
; Sequence 6, Application US/09131247  
; Patent No. 6294170  
; GENERAL INFORMATION:  
; APPLICANT: Boone, Thomas C.  
; APPLICANT: Hershenson, Susan  
; APPLICANT: Bevilacqua, Michael P.  
; APPLICANT: Collins, David S.  
; TITLE OF INVENTION: COMPOSITION AND METHOD FOR TREATING INFLAMMATORY  
; DISEASES  
; FILE REFERENCE: A-365F

QY 121 KGQPREQVYTLPPSRDELTKNOVSLTCLVKGFPDIAVEWESNGQPENNYKTTTPPVLD 180  
Db 240 KGQPREQVYTLPPSRDELTKNOVSLTCLVKGFPDIAVEWESNGQPENNYKTTTPPVLD 299  
QY 181 SDGSFFLSKLTVDKSRWQGNVFCVSMHEALHNNHYTKLSLSPGK 228  
Db 300 SDGSFFLSKLTVDKSRWQGNVFCVSMHEALHNNHYTKLSLSPGK 347

## RESULT 6

US-08-466-465-8  
; Sequence 8, Application US/08466465  
; Patent No. 6162432  
; GENERAL INFORMATION:  
; APPLICANT: Wallner, Barbara P.  
; APPLICANT: Cooper, Kevin D.  
; TITLE OF INVENTION: Method of Prophylaxis or Treatment of Antigen  
; TITLE OF INVENTION: Presenting Cell Driven Skin Conditions Using  
; TITLE OF INVENTION: Inhibitors of the CD2/LFA-3 Interaction  
; NUMBER OF SEQUENCES: 8  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: LAHIVE & COCKFIELD  
; STREET: 60 State Street, Suite 510  
; CITY: Boston  
; STATE: Massachusetts  
; COUNTRY: USA  
; ZIP: 02109-1875  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patentin Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/466.465  
; FILING DATE:  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: PCT/US92/08755  
; FILING DATE: 06-OCT-1992  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 07/862.022  
; FILING DATE: 12-APR-1992  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 07/770.969  
; FILING DATE: 07-OCT-1991  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Myers, Louis (PLM)  
; REGISTRATION NUMBER: 35,965  
; REFERENCE/DOCKET NUMBER: BGP-111CP  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (617)227-7400  
; TELEFAX: (617)227-5941  
; INFORMATION FOR SEQ ID NO: 8:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 347 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
; US-08-466-465-8

Query Match 92.5%; Score 1234; DB 4; Length 347;  
Best Local Similarity 99.6%; Pred. No. 1.le-42;  
Matches 227; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 MDKTHTCPPCAPPELLGGPSVFLFPPPKDITLMISRTPEVTCVVVDVSHEDPEVKFNWYV 60  
Db 120 VDKTHTCPPCAPPELLGGPSVFLFPPPKDITLMISRTPEVTCVVVDVSHEDPEVKFNWYV 179  
QY 61 DGVEVHNAKTKPREEQYNSTYRVVSVLTVLHQDWLNGKEYKCKVSNKALPAPIEKTISKA 120  
Db 180 DGVEVHNAKTKPREEQYNSTYRVVSVLTVLHQDWLNGKEYKCKVSNKALPAPIEKTISKA 239  
QY 121 KGQPREQVYTLPPSRDELTKNOVSLTCLVKGFPDIAVEWESNGQPENNYKTTTPPVLD 180

Db 240 KGQPREQVYTLPPSRDELTKNOVSLTCLVKGFPDIAVEWESNGQPENNYKTTTPPVLD 299  
QY 181 SDGSFFLSKLTVDKSRWQGNVFCVSMHEALHNNHYTKLSLSPGK 228  
Db 300 SDGSFFLSKLTVDKSRWQGNVFCVSMHEALHNNHYTKLSLSPGK 347

## RESULT 7

PCT-US92-02050-43  
; Sequence 43, Application PC/TUS9202050  
; GENERAL INFORMATION:  
; APPLICANT: BIOGEN, INC.  
; APPLICANT: WALLNER, Barbara P.  
; APPLICANT: MILLER, Glenn T.  
; APPLICANT: ROSA, Margaret D.  
; TITLE OF INVENTION: CD2-BINDING DOMAIN OF LYMPHOCYTE  
; TITLE OF INVENTION: FUNCTION ASSOCIATED ANTIGEN 3  
; NUMBER OF SEQUENCES: 43  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Fish & Neave  
; STREET: 875 Third Avenue  
; CITY: New York  
; STATE: New York  
; COUNTRY: U.S.A.  
; ZIP: 10022-6250  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patentin Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: PCT/US92/02050  
; FILING DATE: 19920312  
; CLASSIFICATION: 435  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 07/667,971  
; FILING DATE: 12-MAR-1991  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 07/770,967  
; FILING DATE: 07-OCT-1991  
; ATTORNEY/AGENT INFORMATION:  
; NAME: HALEY, James F., Jr.  
; REGISTRATION NUMBER: 27,794  
; REFERENCE/DOCKET NUMBER: B151CIP2  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (212)715-0600  
; TELEFAX: (212)715-0673  
; TELEX: 14-8367  
; INFORMATION FOR SEQ ID NO: 43:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 347 amino acids  
; TYPE: AMINO ACID  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
; PCT-US92-02050-43

Query Match 92.5%; Score 1234; DB 5; Length 347;  
Best Local Similarity 99.6%; Pred. No. 1.le-42;  
Matches 227; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 MDKTHTCPPCAPPELLGGPSVFLFPPPKDITLMISRTPEVTCVVVDVSHEDPEVKFNWYV 60  
Db 120 VDKTHTCPPCAPPELLGGPSVFLFPPPKDITLMISRTPEVTCVVVDVSHEDPEVKFNWYV 179  
QY 61 DGVEVHNAKTKPREEQYNSTYRVVSVLTVLHQDWLNGKEYKCKVSNKALPAPIEKTISKA 120  
Db 180 DGVEVHNAKTKPREEQYNSTYRVVSVLTVLHQDWLNGKEYKCKVSNKALPAPIEKTISKA 239  
QY 121 KGQPREQVYTLPPSRDELTKNOVSLTCLVKGFPDIAVEWESNGQPENNYKTTTPPVLD 180  
Db 240 KGQPREQVYTLPPSRDELTKNOVSLTCLVKGFPDIAVEWESNGQPENNYKTTTPPVLD 299  
QY 181 SDGSFFLSKLTVDKSRWQGNVFCVSMHEALHNNHYTKLSLSPGK 228

QY 181 SDGSFFLYSKLTVDKSRWQGNVFCVSMHEALHNHYTKQSLSPGK 228  
|||||  
Db 300 SDGSFFLYSKLTVDKSRWQGNVFCVSMHEALHNHYTKQSLSPGK 347

## RESULT 4

US-08-459-657-43  
; Sequence 43, Application US/08459657  
; Patent No. 5914111  
; GENERAL INFORMATION:  
; APPLICANT: BIOGEN, INC.  
; APPLICANT: WALLNER, Barbara P.  
; APPLICANT: MILLER, Glenn T.  
; APPLICANT: ROSA, Margaret D.  
; TITLE OF INVENTION: CD2-BINDING DOMAIN OF LYMPHOCYTE  
; TITLE OF INVENTION: FUNCTION ASSOCIATED ANTIGEN 3  
; NUMBER OF SEQUENCES: 43  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Fish & Neave  
; STREET: 875 Third Avenue  
; CITY: New York  
; STATE: New York  
; COUNTRY: U.S.A.  
; ZIP: 10022-6250  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent In Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/459,657  
; FILING DATE: 02-JUN-1995  
; CLASSIFICATION: 424  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: PCT/US92/02050  
; FILING DATE: 12-MAR-1992  
; APPLICATION NUMBER: US 07/667,971  
; FILING DATE: 12-MAR-1991  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 07/770,967  
; FILING DATE: 07-OCT-1991  
; ATTORNEY/AGENT INFORMATION:  
; NAME: HALEY, James F., Jr.  
; REGISTRATION NUMBER: 27,794  
; REFERENCE/DOCKET NUMBER: B151CIP2  
; TELEPHONE: (212)715-0600  
; TELEFAX: (212)715-0673  
; TELEX: 14-8367  
; INFORMATION FOR SEQ ID NO: 43:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 347 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
US-08-459-657-43

Query Match 92.5%; Score 1234; DB 2; Length 347;  
Best Local Similarity 99.6%; Pred. No. 1.1e-42;  
Matches 227; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 MDKTHCPCPAPELLGGPSVFLFPPKPKDTLMISRTPEVTCVVDVSHEDPEVKFNWYV 60  
:|||||  
Db 120 VDKTHCPCPAPELLGGPSVFLFPPKPKDTLMISRTPEVTCVVDVSHEDPEVKFNWYV 179  
|||||  
QY 61 DGVEVHNATKPREQYNSYRVSVLTVLHODWLNKGEYKCKVSNKALPAPIEKTISKA 120  
|||||  
Db 180 DGVEVHNATKPREQYNSYRVSVLTVLHODWLNKGEYKCKVSNKALPAPIEKTISKA 239  
|||||  
QY 121 KGQPREQVYTLPPSRDELTKNQVSLTCLVKGFYPSDIAVEGWSNQENPKYKTTTPPVLD 180  
|||||  
Db 240 KGQPREQVYTLPPSRDELTKNQVSLTCLVKGFYPSDIAVEGWSNQENPKYKTTTPPVLD 299  
|||||

QY 181 SDGSFFLYSKLTVDKSRWQGNVFCVSMHEALHNHYTKQSLSPGK 228  
|||||  
Db 300 SDGSFFLYSKLTVDKSRWQGNVFCVSMHEALHNHYTKQSLSPGK 347

## RESULT 5

US-08-460-132-43  
; Sequence 43, Application US/08460132  
; Patent No. 5928643  
; GENERAL INFORMATION:  
; APPLICANT: BIOGEN, INC.  
; APPLICANT: WALLNER, Barbara P.  
; APPLICANT: MILLER, Glenn T.  
; APPLICANT: ROSA, Margaret D.  
; TITLE OF INVENTION: CD2-BINDING DOMAIN OF LYMPHOCYTE  
; TITLE OF INVENTION: FUNCTION ASSOCIATED ANTIGEN 3  
; NUMBER OF SEQUENCES: 43  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Fish & Neave  
; STREET: 875 Third Avenue  
; CITY: New York  
; STATE: New York  
; COUNTRY: U.S.A.  
; ZIP: 10022-6250  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent In Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/460,132  
; FILING DATE: 02-JUN-1995  
; CLASSIFICATION: 424  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 07/940,861  
; FILING DATE: 21-OCT-1992  
; APPLICATION NUMBER: PCT/US92/02050  
; FILING DATE: 12-MAR-1992  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 07/667,971  
; FILING DATE: 12-MAR-1991  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 07/770,967  
; FILING DATE: 07-OCT-1991  
; ATTORNEY/AGENT INFORMATION:  
; NAME: HALEY, James F., Jr.  
; REGISTRATION NUMBER: 27,794  
; REFERENCE/DOCKET NUMBER: B151CIP2  
; TELEPHONE: (212)715-0600  
; TELEFAX: (212)715-0673  
; TELEX: 14-8367  
; INFORMATION FOR SEQ ID NO: 43:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 347 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
US-08-460-132-43

Query Match 92.5%; Score 1234; DB 2; Length 347;  
Best Local Similarity 99.6%; Pred. No. 1.1e-42;  
Matches 227; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 MDKTHCPCPAPELLGGPSVFLFPPKPKDTLMISRTPEVTCVVDVSHEDPEVKFNWYV 60  
:|||||  
Db 120 VDKTHCPCPAPELLGGPSVFLFPPKPKDTLMISRTPEVTCVVDVSHEDPEVKFNWYV 179  
|||||  
QY 61 DGVEVHNATKPREQYNSYRVSVLTVLHODWLNKGEYKCKVSNKALPAPIEKTISKA 120  
|||||  
Db 180 DGVEVHNATKPREQYNSYRVSVLTVLHODWLNKGEYKCKVSNKALPAPIEKTISKA 239  
|||||

QY 239 -----AEKLKE 244  
DB 310 GDETRQLQLEAVNITDLSENKQ 331

RESULT 2  
US-07-940-861-43  
; Sequence 43, Application US/07940861  
; Patent No. 5547853  
; GENERAL INFORMATION:  
; APPLICANT: BIOGEN, INC.  
; APPLICANT: WALLNER, Barbara P.  
; APPLICANT: MILLER, Glenn T.  
; APPLICANT: ROSA, Margaret D.  
; TITLE OF INVENTION: CD2-BINDING DOMAIN OF LYMPHOCYTE  
; TITLE OF INVENTION: FUNCTION ASSOCIATED ANTIGEN 3  
; NUMBER OF SEQUENCES: 43  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Fish & Neave  
; STREET: 875 Third Avenue  
; CITY: New York  
; STATE: New York  
; COUNTRY: U.S.A.  
; ZIP: 10022-6250  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent In Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/07/940,861  
; FILING DATE: 21-OCT-1992  
; CLASSIFICATION: 514  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: PCT/US92/02050  
; FILING DATE: 12-MAR-1992  
; APPLICATION NUMBER: US 07/667,971  
; FILING DATE: 12-MAR-1991  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 07/770,967  
; FILING DATE: 07-OCT-1991  
; ATTORNEY/AGENT INFORMATION:  
; NAME: HALEY, James F., Jr.  
; REGISTRATION NUMBER: 27,794  
; REFERENCE/DOCKET NUMBER: B151CIP2  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (212)715-0600  
; TELEFAX: (212)715-0673  
; TELEX: 14-8367  
; INFORMATION FOR SEQ ID NO: 43:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 347 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
US-07-940-861-43

Query Match 92.5%; Score 1234; DB 1; Length 347;  
Best Local Similarity 99.6%; Pred. No. 1.1e-42;  
Matches 227; Conservative 1; Mismatches 0; Indels 0; Gaps 0;  
QY 1 MDKTHCTPCPAPELLGGPSVFLFPPKPKDTLMISRTPEVTCVVDVSHEDPEVKFNWYV 60  
DB 120 VDKTHCTPCPAPELLGGPSVFLFPPKPKDTLMISRTPEVTCVVDVSHEDPEVKFNWYV 179  
QY 61 DGVEVHNATKPREQYNTYRVSVLTVLHQDLNGKEYCKVSNKALPAPIETKISK 120  
DB 180 DGVEVHNATKPREQYNTYRVSVLTVLHQDLNGKEYCKVSNKALPAPIETKISK 239  
QY 121 KGQPREPQVYTLPPSRDELTKNOVSLTCLVKGFPSPDI AVEWESNGQPNKYKTTTPPVL 180  
DB 240 KGQPREPQVYTLPPSRDELTKNOVSLTCLVKGFPSPDI AVEWESNGQPNKYKTTTPPVL 299

QY 181 SDGSFFLYSKLTVDKSRWQGNVFCSCVMHEALHNHYTKSLSPGK 228  
DB 300 SDGSFFLYSKLTVDKSRWQGNVFCSCVMHEALHNHYTKSLSPGK 347

RESULT 3  
US-08-459-512-43  
; Sequence 43, Application US/08459512  
; Patent No. 5728677  
; GENERAL INFORMATION:  
; APPLICANT: BIOGEN, INC.  
; APPLICANT: WALLNER, Barbara P.  
; APPLICANT: MILLER, Glenn T.  
; APPLICANT: ROSA, Margaret D.  
; TITLE OF INVENTION: CD2-BINDING DOMAIN OF LYMPHOCYTE  
; TITLE OF INVENTION: FUNCTION ASSOCIATED ANTIGEN 3  
; NUMBER OF SEQUENCES: 43  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Fish & Neave  
; STREET: 875 Third Avenue  
; CITY: New York  
; STATE: New York  
; COUNTRY: U.S.A.  
; ZIP: 10022-6250  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent In Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/459,512  
; FILING DATE: 02-JUN-1995  
; CLASSIFICATION: 514  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: PCT/US92/02050  
; FILING DATE: 12-MAR-1992  
; APPLICATION NUMBER: US 07/667,971  
; FILING DATE: 12-MAR-1991  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 07/770,967  
; FILING DATE: 07-OCT-1991  
; ATTORNEY/AGENT INFORMATION:  
; NAME: HALEY, James F., Jr.  
; REGISTRATION NUMBER: 27,794  
; REFERENCE/DOCKET NUMBER: B151CIP2  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (212)715-0600  
; TELEFAX: (212)715-0673  
; TELEX: 14-8367  
; INFORMATION FOR SEQ ID NO: 43:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 347 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
US-08-459-512-43

Query Match 92.5%; Score 1234; DB 1; Length 347;  
Best Local Similarity 99.6%; Pred. No. 1.1e-42;  
Matches 227; Conservative 1; Mismatches 0; Indels 0; Gaps 0;  
QY 1 MDKTHCTPCPAPELLGGPSVFLFPPKPKDTLMISRTPEVTCVVDVSHEDPEVKFNWYV 60  
DB 120 VDKTHCTPCPAPELLGGPSVFLFPPKPKDTLMISRTPEVTCVVDVSHEDPEVKFNWYV 179  
QY 61 DGVEVHNATKPREQYNTYRVSVLTVLHQDLNGKEYCKVSNKALPAPIETKISK 120  
DB 180 DGVEVHNATKPREQYNTYRVSVLTVLHQDLNGKEYCKVSNKALPAPIETKISK 239  
QY 121 KGQPREPQVYTLPPSRDELTKNOVSLTCLVKGFPSPDI AVEWESNGQPNKYKTTTPPVL 180  
DB 240 KGQPREPQVYTLPPSRDELTKNOVSLTCLVKGFPSPDI AVEWESNGQPNKYKTTTPPVL 299

GenCore version 5.1.4.p5.4578  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: April 21, 2003, 10:42:26 ; Search time 13.7471 Seconds  
(without alignments)  
526.515 Million cell updates/sec

Title: 2LINK7

Perfect score: 1334

Sequence: 1 MDKTHCPCPAPELLGGPS.....GRDLKAFYDKVAELKEAF 246

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.0

Searched: 262574 seqs, 29422922 residues

Total number of hits satisfying chosen parameters: 262574

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued Patents.AA.\*

- 1: /cgn2.6/ptodata/1/iaa/5A\_COMB.pep.\*
- 2: /cgn2.6/ptodata/1/iaa/5B\_COMB.pep.\*
- 3: /cgn2.6/ptodata/1/iaa/6A\_COMB.pep.\*
- 4: /cgn2.6/ptodata/1/iaa/6B\_COMB.pep.\*
- 5: /cgn2.6/ptodata/1/iaa/PCTUS\_COMB.pep.\*
- 6: /cgn2.6/ptodata/1/iaa/backfiles1.pep.\*

pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1238	92.8	389	4	US-09-131-247-14
2	1234	92.5	347	1	US-07-940-861-43
3	1234	92.5	347	1	US-08-459-512-43
4	1234	92.5	347	2	US-08-459-657-43
5	1234	92.5	347	2	US-08-460-132-43
6	1234	92.5	347	4	US-08-466-465-8
7	1234	92.5	347	5	PCT-US92-02050-43
8	1234	92.5	482	4	US-09-189-129-2
9	1233	92.4	232	2	US-08-595-043A-50
10	1233	92.4	235	4	US-09-131-247-6
11	1233	92.4	331	4	US-09-178-869-2
12	1233	92.4	360	4	US-09-180-100-11
13	1233	92.4	371	1	US-08-236-311-7
14	1233	92.4	371	3	US-08-457-918-7
15	1233	92.4	376	4	US-09-180-100-22
16	1233	92.4	387	1	US-08-470-299-4
17	1233	92.4	388	4	US-09-131-247-16
18	1233	92.4	396	2	US-08-784-512-3
19	1233	92.4	396	4	US-09-176-228-3
20	1233	92.4	424	4	US-09-333-593A-8
21	1233	92.4	424	5	PCT-US95-03866-12
22	1233	92.4	424	5	PCT-US95-03866-14
23	1233	92.4	437	5	PCT-US96-10043-11
24	1233	92.4	442	5	PCT-US96-10043-9
25	1233	92.4	446	3	US-08-397-411-7
26	1233	92.4	449	1	US-08-458-516-13
27	1233	92.4	459	1	US-08-157-101A-7

ALIGNMENTS

RESULT 1

US-09-131-247-14  
; Sequence 14, Application US/09131247

; Patent No. 6294170

; GENERAL INFORMATION:

; APPLICANT: Boone, Thomas C.

; APPLICANT: Hersenson, Susan

; APPLICANT: Bevilacqua, Michael P.

; APPLICANT: Collins, David S.

; TITLE OF INVENTION: COMPOSITION AND METHOD FOR TREATING INFLAMMATORY

; TITLE OF INVENTION: DISEASES

; FILE REFERENCE: A-365F

; CURRENT APPLICATION NUMBER: US/09/131.247

; EARLIER FILING DATE: 1998-08-07

; EARLIER APPLICATION NUMBER: 60/055,185

; EARLIER FILING DATE: 1997-08-08

; EARLIER APPLICATION NUMBER: PCT/US 97/02131

; EARLIER FILING DATE: 1997-02-10

; NUMBER OF SEQ ID NOS: 16

; SOFTWARE: PatentIn Ver. 2.0

; SEQ ID NO 14

; LENGTH: 389

; TYPE: PRT

; ORGANISM: Human

US-09-131-247-14

Query Match

Best Local Similarity 92.8%; Score 1238; DB 4; Length 389;

Matches 234; Conservative 4; Mismatches 5; Indels 79; Gaps 4;

QY	2	DKTHTCPCPAPELLGGPSVFLEPPPKDPLMSRTPEVTCVVVDVSHEDPEVKFNMYVD	61
Db	10	DKTHTCPCPAPELLGGPSVFLEPPPKDPLMSRTPEVTCVVVDVSHEDPEVKFNMYVD	69
QY	62	GVEVHNAKTPREEQNSTYRVVSVLTVLHQDWLNGKEYCKVSNKALPAPIEKTISKAK	121
Db	70	GVEVHNAKTPREEQNSTYRVVSVLTVLHQDWLNGKEYCKVSNKALPAPIEKTISKAK	129
QY	122	GQPREQVYTLPPSRDELTKNQVSLTCLVKGFYPSDIAVEWESNGQPENNYKTTTPPVLD	181
Db	130	GQPREQVYTLPPSRDELTKNQVSLTCLVKGFYPSDIAVEWESNGQPENNYKTTTPPVLD	189
QY	182	DGSEFFLYSKLTVDKSRWQGNVFSCVSMHEALHNHYTQKSLSLSPGK-----	228
Db	190	DGSEFFLYSKLTVDKSRWQGNVFSCVSMHEALHNHYTQKSLSLSPGKMRPSGRKSKMQA	249
QY	229	--DW---LKAFY-----DKV-----	238
Db	250	FRWDVYNQKTFYLRNQLVAGYLGQPNVNLEEKIDVYVPIEPHALFLGHHGKMKLSCVKS	309







KW cancer; hepatitis; acquired immunodeficiency syndrome; AIDS;  
KW autoimmune disorder; transplant rejection; cardiovascular disease;  
KW arteriosclerosis; Fc-huAGP-1; fusion protein.

OS Homo sapiens.

XX WO2000063253-A1.

XX 26-OCT-2000.

XX 24-MAR-2000; 2000WO-US08004.

XX 16-APR-1999; 99US-0293245.

XX (AMGE-) AMGEN INC.

XX Hsu H, Meng S;

XX WPI; 2000-665240/64.

XX N-PSDB; AAC67833.

XX Fusion protein of AGP-1 protein and an Fc region, used to treat  
PT proliferative disorders, immune disorders, and virally-induced  
PT disorders -

XX Disclosure; Fig 4; 93pp; English.

XX The present sequence is an AGP-1 fusion protein. AGP-1 is a  
QC type II transmembrane protein. The fusion proteins comprise an Fc  
QC immunoglobulin region fused to the N-terminal portion of the AGP-1  
QC protein. The fusion proteins can be used to induce apoptosis in a tissue,  
QC and to treat proliferative disorders, immune disorders, or  
QC virally-induced disorders. The proliferative disorders include cancers,  
QC such as breast, prostate, lung or colon cancer. The viral infections  
QC include hepatitis, and acquired immunodeficiency syndrome (AIDS), and the  
QC immune disorders may be autoimmune disorders or transplant rejection.  
QC Cardiovascular diseases such as arteriosclerosis may also be treated. The  
QC AGP-1 containing fusion proteins have increased biological activity  
QC compared to the soluble AGP-1 proteins used in prior art therapies.

XX Sequence 423 AA;

XX Query Match 93.3%; Score 1245; DB 21; Length 423;  
XX Best Local Similarity 64.8%; Pred. No. 6e-35;  
XX Matches 237; Conservative 1; Mismatches 5; Indels 123; Gaps 4;

Qy 2 DKHTCCPCAPPELLGGPSVFLFPPKPKDTLMISRTPEVTCVVDVSHEDPEVKFNWYVD 61

Db 29 DKHTCCPCAPPELLGGPSVFLFPPKPKDTLMISRTPEVTCVVDVSHEDPEVKFNWYVD 88

Qy 62 GVEVHNATKPREQYNSTYRVVSVLTVLHQDWLNGKEYCKVSNKALPAPIETISKAK 121

Db 89 GVEVHNATKPREQYNSTYRVVSVLTVLHQDWLNGKEYCKVSNKALPAPIETISKAK 148

Qy 122 GQPREPQVYTLPPSDELTKNOVSLTCLVKGFYPSDIAVEWESNGQPENNYKTPPVLD 181

Db 149 GQPREPQVYTLPPSDELTKNOVSLTCLVKGFYPSDIAVEWESNGQPENNYKTPPVLD 208

Qy 182 DGSFPLYSLKLVDRSRQGNVFCVSMHEALHNYTKSLSPGK----- 228

Db 209 DGSFPLYSLKLVDRSRQGNVFCVSMHEALHNYTKSLSPGKVRERQPVRAHI 268

Qy 229 -----DW-----LKAFY----- 235

Db 269 TGTGRSNTLSPNSKNEKALGRKINSWESSRSGHSFLSLHNLRLNGELVIHEKGYIYS 328

Qy 236 -----DKVAE----- 240

Db 329 QTYFRQBEIKENTKNDKQVQYIYKTSYTPDPILLMKASARNCSWSDAEYGLYSIQGG 388

Qy 241 --KLKE 244

Db 389 IFELKE 394

RESULT 15

AAB28692

ID AAB28692 standard; Protein; 441 AA.

XX AAB28692;

XX 14-FEB-2001 (first entry)

XX Fc-huAGP-1 (95-281) fusion protein.

XX Human; AGP-1; type II transmembrane protein; cytostatic; antiviral;  
KW antiinflammatory; hepatotropic; antiarteriosclerotic; anti-HIV; HIV;  
KW human immunodeficiency virus; apoptosis; proliferative disorder;  
KW cancer; hepatitis; acquired immunodeficiency syndrome; AIDS;  
KW autoimmune disorder; transplant rejection; cardiovascular disease;  
KW arteriosclerosis; Fc-huAGP-1; fusion protein.

XX Homo sapiens.

XX WO2000063253-A1.

XX 26-OCT-2000.

XX 24-MAR-2000; 2000WO-US08004.

XX 16-APR-1999; 99US-0293245.

XX (AMGE-) AMGEN INC.

XX Hsu H, Meng S;

XX WPI; 2000-665240/64.

XX N-PSDB; AAC67832.

XX Fusion protein of AGP-1 protein and an Fc region, used to treat  
PT proliferative disorders, immune disorders, and virally-induced  
PT disorders -

XX Disclosure; Fig 3; 93pp; English.

XX The present sequence is an AGP-1 fusion protein. AGP-1 is a  
CC type II transmembrane protein. The fusion proteins comprise an Fc  
CC immunoglobulin region fused to the N-terminal portion of the AGP-1  
CC protein. The fusion proteins can be used to induce apoptosis in a tissue,  
CC and to treat proliferative disorders, immune disorders, or  
CC virally-induced disorders. The proliferative disorders include cancers,  
CC such as breast, prostate, lung or colon cancer. The viral infections  
CC include hepatitis, and acquired immunodeficiency syndrome (AIDS), and the  
CC immune disorders may be autoimmune disorders or transplant rejection.  
CC Cardiovascular diseases such as arteriosclerosis may also be treated. The  
CC AGP-1 containing fusion proteins have increased biological activity  
CC compared to the soluble AGP-1 proteins used in prior art therapies.

XX Sequence 441 AA;

XX Query Match 93.3%; Score 1245; DB 21; Length 441;  
XX Best Local Similarity 61.6%; Pred. No. 6.6e-35;  
XX Matches 237; Conservative 1; Mismatches 5; Indels 142; Gaps 4;

Qy 2 DKHTCCPCAPPELLGGPSVFLFPPKPKDTLMISRTPEVTCVVDVSHEDPEVKFNWYVD 61

Db 29 DKHTCCPCAPPELLGGPSVFLFPPKPKDTLMISRTPEVTCVVDVSHEDPEVKFNWYVD 88

Qy 62 GVEVHNATKPREQYNSTYRVVSVLTVLHQDWLNGKEYCKVSNKALPAPIETISKAK 121

Db 89 GVEVHNATKPREQYNSTYRVVSVLTVLHQDWLNGKEYCKVSNKALPAPIETISKAK 148

Qy 122 GQPREPQVYTLPPSDELTKNOVSLTCLVKGFYPSDIAVEWESNGQPENNYKTPPVLD 181

Db 149 GQPREPQVYTLPPSDELTKNOVSLTCLVKGFYPSDIAVEWESNGQPENNYKTPPVLD 208

Db 121 KGQPREPQVYTLPPSRDELTKNQVSLTCLVKGFYPSDIAVEWESNGQPENNYKTTTPVLD 180  
QY 181 SDGSFPLYSLTVDKSRWQGNVFCVSMHEALHNHYTOKSLSPGKD----- 229  
Db 181 SDGSFPLYSLTVDKSRWQGNVFCVSMHEALHNHYTOKSLSPGKETFPKYLHYDE 240  
QY 230 -----WLK----- 232  
Db 241 ETSHQLLCKDCKPPGTLYLKQHTAKWTKVACPCPDHYTDSWHTSDCLCYCSPVKELQYV 300  
QY 233 ----- 232  
Db 301 KQECNRTHNRVCEKGEGRYLEIEFCLKHSRCPGPGFVGWQAGTPERTNVCKRCPDGFFSNE 360  
QY 233 -----AFYDKVAEKLKEA 245  
Db 361 TSSKAPCRKHTNCVFGLLLTQGNATHDNCISGNS 398  
RESULT 13  
AAB80904  
ID AAB80904 standard; protein; 401 AA.  
XX AAB80904;  
XX  
XX 31-MAY-2001 (first entry)  
XX Human metFcdeltaC-OPG(22-194) fusion protein.  
XX  
KW Human; anticancer; Antimetastatic; Osteogenic; lytic bone disease;  
KW multiple myeloma; osteosclerotic bone metastasis; OPG; osteoprotegrin;  
KW osteoclast formation inhibition; bone resorption inhibition;  
KW immunoglobulin.  
XX  
OS Homo sapiens.  
XX  
PN WO200117543-A2.  
XX  
PD 15-MAR-2001.  
PF 18-AUG-2000; 2000WO-US22806.  
XX  
PR 03-SEP-1999; 99US-0389545.  
XX  
PA (AMGE-) AMGEN INC.  
XX  
PI Dunstan CR;  
XX  
DR WPI; 2001-265936/27.  
XX  
PT Preventing or treating lytic bone diseases, particularly associated  
PT with cancer or metastasis, by administering an osteoprotegrin  
XX polypeptide  
XX  
PS Claim 11; Fig 8; 87pp; English.  
XX  
CC The present invention relates to a method for the prevention or treatment  
CC of lytic bone disease or multiple myeloma. Also the method can be used  
CC for preventing metastasis of cancer to bone or osteosclerotic bone  
CC metastasis. The method comprises administering an OPG (osteoprotegrin)  
CC polypeptide or OPG fusion protein. The present sequence is one such OPG  
CC fusion protein. OPG inhibits formation of osteoclasts (and thus bone  
CC resorption) by blocking differentiation from monocytes/macrophage  
CC precursors. The OPG polypeptide can be used in a method of preventing or  
CC treating lytic bone disease, for preventing metastasis of cancer to bone  
CC (e.g. breast, prostate, thyroid, cancer of the kidney, lung, oesophogeal,  
CC rectal, bladder, cervical, ovarian, liver, cancer of the gastrointestinal  
CC tract, multiple myeloma or lymphoma) and preventing the osteosclerotic  
CC bone metastasis. The OPG fusion polypeptides are used in the prevention  
CC or treatment of loss of bone mass, which occurs in conditions including  
CC osteoporosis, such as primary osteoporosis, endocrine osteoporosis  
CC (hyperparathyroidism, Cushing's syndrome and acromegaly), hereditary and

congenital forms of osteoporosis (osteogenesis imperfecta,  
homocystinuria, Menkes' syndrome and Riley-Day syndrome) and osteoporosis  
due to immobilisation of extremities; Paget's disease of bone (osteitis  
deformans) in adults and juveniles; osteomyelitis, or an infectious  
lesion in bone; hypercalcaemia resulting from solid tumours (breast, lung  
and kidney) and haematologic malignancies (multiple myeloma, lymphoma and  
leukaemia), idiopathic hypercalcaemia, and hypercalcaemia associated with  
hyperthyroidism and renal function disorders; osteopaenia following  
surgery, induced by steroid administration, and associated with disorders  
of the small and large intestine and with chronic hepatic and renal  
diseases; osteonecrosis, or bone cell death, associated with traumatic  
injury or nontraumatic necrosis associated with Gaucher's disease, sickle  
cell anaemia, systemic lupus erythematosus and other conditions; bone  
loss due to rheumatoid arthritis; periodontal bone loss; osteoarthritis;  
prosthetic loosening, and osteolytic metastasis. The OPG fusion proteins  
are also used in the replacement of structurally sound bone with  
disorganised bone as seen in Paget's disease of bone (osteitis deformans)  
in adults and juveniles; hyperparathyroidism, in congenital bone  
disorders such as fibrous dysplasia, and in osteosclerotic bone  
metastases. The OPG fusion proteins can exhibit increased circulating  
half-lives and slower clearance times, thereby providing a more sustained  
activity. The OPG fusion protein comprises a fragment of the human OPG  
protein and the Fc region of immunoglobulin IgGgamma1 (the hinge, CH2 and  
CH3 regions; see AAB80897-8).  
XX  
XX  
SQ Sequence 401 AA;  
Query Match 93.3%; Score 1245; DB 22; Length 401;  
Best Local Similarity 58.5%; Pred. No. 5.3e-35;  
Matches 233; Conservative 5; Mismatches 7; Indels 153; Gaps 2;  
QY 1 MKDTHTCPCPAPELGGPSVLFPPKPKYDITLMISRTPEVTCVVVDVSHEDPEVKFNWYV 60  
Db 1 MKDTHTCPCPAPELGGPSVLFPPKPKYDITLMISRTPEVTCVVVDVSHEDPEVKFNWYV 60  
QY 61 DGVEVHNATKPREEQNSTYRVSVTLVTHQDNLNGKEYCKVSKNKAALPAPIEKTISKA 120  
Db 61 DGVEVHNATKPREEQNSTYRVSVTLVTHQDNLNGKEYCKVSKNKAALPAPIEKTISKA 120  
QY 121 KGQPREPQVYTLPPSRDELTKNQVSLTCLVKGFYPSDIAVEWESNGQPENNYKTTTPVLD 180  
Db 121 KGQPREPQVYTLPPSRDELTKNQVSLTCLVKGFYPSDIAVEWESNGQPENNYKTTTPVLD 180  
QY 181 SDGSFPLYSLTVDKSRWQGNVFCVSMHEALHNHYTOKSLSPGKD----- 229  
Db 181 SDGSFPLYSLTVDKSRWQGNVFCVSMHEALHNHYTOKSLSPGKETFPKYLHYDE 240  
QY 230 -----WLK----- 232  
Db 241 ETSHQLLCKDCKPPGTLYLKQHTAKWTKVACPCPDHYTDSWHTSDCLCYCSPVKELQYV 300  
QY 233 ----- 232  
Db 301 KQECNRTHNRVCEKGEGRYLEIEFCLKHSRCPGPGFVGWQAGTPERTNVCKRCPDGFFSNE 360  
QY 233 -----AFYDKVAEKLKEA 245  
Db 361 TSSKAPCRKHTNCVFGLLLTQGNATHDNCISGNS 398  
RESULT 14  
AAB28693  
ID AAB28693 standard; Protein; 423 AA.  
XX AAB28693;  
XX  
XX 14-FEB-2001 (first entry)  
XX Fc-huAGP-1 (114-281) fusion protein.  
DE Human; AGP-1; type II transmembrane protein; cytostatic; antiviral;  
KW antinflammatory; hepatotropic; antiarteriosclerotic; anti-HIV; HIV;  
KW human immunodeficiency virus; apoptosis; proliferative disorder;

OS Homo sapiens.  
XX Synthetic.  
XX Key Location/Qualifiers  
PH Protein 2..374  
FT Region /note= "Recombinant human Fc-OB variant 2 protein"  
FT Region 229..374  
FT Region /note= "Human OB protein"  
XX  
XX WO9828427-A1.  
XX  
XX PD 02-JUL-1998.  
XX  
XX PF 11-DEC-1997; 97WO-US23183.  
XX PR 20-DEC-1996; 96US-0770973.  
XX  
XX PA (AMGE-) AMGEN INC.  
XX  
XX PI Hecht RI, Mann MB;  
XX  
XX DR WPI: 1998-377658/32.  
XX DR N-PSDB; AAV32902.  
XX  
XX PT New fusion proteins of OB and Fc - used for treating e.g. excess  
PT weight, diabetes, arterial sclerosis, arterial plaque, high blood  
PT lipid level, gall stones or stroke  
XX  
XX PS Claim 2; Fig 5A-5C; 107pp; English.  
XX  
XX CC The present sequence represents a recombinant human MetFc-OB variant 2  
CC fusion protein having a 5 residue deletion of residues 2-6 of the  
CC wild-type Fc-OB protein sequence shown in AAM49073. The invention  
CC provides Fc-OB fusion proteins whereby the Fc region of an immunoglobulin  
CC or its analogue is linked, either directly or indirectly using a linker,  
CC to the N-terminus of an OB protein or its analogue. The Fc-OB fusion  
CC proteins are claimed to demonstrate increased stability and clearance  
CC rate and decreased degradation as compared to OB protein or a fusion of  
CC Fc to the C-terminus of the OB protein. These Fc-OB fusion proteins are  
CC also claimed to be useful for treating excess weight in an individual or  
CC animal or for treating co-morbidities associated with excess fat such as  
CC diabetes, high blood lipid level, arterial sclerosis and stroke.  
XX  
XX SQ Sequence 374 AA;  
  
Query Match 93.3%; Score 1245; DB 19; Length 374;  
Best Local Similarity 74.4%; Pred. No. 4.5e-35;  
Matches 233; Conservative 4; Mismatches 7; Indels 59; Gaps 2;  
  
QY 1 MDKTHCTCPAPPELLGGPSVFLFPPKPKDTLMISRTPEVTCVVDVSHEDPEVKFNWYV 60  
DB 1 MDKTHCTCPAPPELLGGPSVFLFPPKPKDTLMISRTPEVTCVVDVSHEDPEVKFNWYV 60  
  
QY 61 DGVEVHNAKTKPREEQNSTYRVWSVLTVLHODWLNKGEYCKVSNKALPAPIETISK 120  
DB 61 DGVEVHNAKTKPREEQNSTYRVWSVLTVLHODWLNKGEYCKVSNKALPAPIETISK 120  
  
QY 121 KGQPREQVYTLPPSRDELTKNQVSLTCLVKGFYPSDIAVEWESNGQPENNYKTPPVLD 180  
DB 121 KGQPREQVYTLPPSRDELTKNQVSLTCLVKGFYPSDIAVEWESNGQPENNYKTPPVLD 180  
  
QY 181 SDGSFFLSKLTVDKSRWQQGNVFCQSVNHEALHNYHTQKSLSLSPGK 228  
DB 181 SDGSFFLSKLTVDKSRWQQGNVFCQSVNHEALHNYHTQKSLSLSPGK 240  
  
QY 229 -----DWLKAFFDKV----- 238  
DB 241 LIKTIIVTRINDISHTQSVSSKOKVTGLDFIPGLHILPILTSKMDQTLAVYQVITSMPSRN 300  
  
QY 239 -----AEKLKE 244  
DB 301 VIQISNDLENLRD 313

RESULT 12  
AAV72922  
ID AAV72922 standard; Protein; 401 AA.  
XX  
AC AAV72922;  
XX  
DT 13-JUN-2001 (first entry)  
XX  
XX Human met-Fc (lacking 1-5 residues)-OPG (22-194 aa) fusion protein.  
XX  
XX Human; fusion protein; osteoprotegerin; OPG; Fc protein; osteopathic;  
KW therapy; bone loss; osteoporosis; Paget's disease; osteomyelitis;  
KW hypercalcaemia; osteopenia; osteonecrosis; rheumatoid arthritis;  
KW osteolytic metastasis; prosthetic loosening; immunoglobulin G1;  
KW IgG1; periodontal.  
XX  
XX OS Homo sapiens.  
XX  
XX Key Location/Qualifiers  
FH Region 229..401  
FT Region /note= "Derived from human osteoprotegerin fragment  
FT (22-194 residues)"  
FT Region 1..228  
FT Region /note= "Met-human IgG1 Fc region lacking 1-5 residues;  
FT Corresponds to 6-231 residues of human IgG1 Fc region"  
XX  
XX PN WO200118203-A1.  
XX  
XX PD 15-MAR-2001.  
XX  
XX PF 18-AUG-2000; 2000WO-US22797.  
XX  
XX PR 03-SEP-1999; 99US-0389782.  
XX  
XX PA (AMGE-) AMGEN INC.  
XX  
XX PI Dunstan CR, Wooden SK, Mann MB;  
XX  
XX DR WPI: 2001-244572/25.  
XX  
XX PT Osteoprotegerin-Fc protein fusions useful for treating bone loss caused  
PT by e.g. osteoporosis, Paget's disease and osteomyelitis -  
XX  
XX PS Claim 7; Fig 8; 119pp; English.  
XX  
XX CC The present sequence is a fusion protein comprising human met-Fc region  
CC (lacking 1-5 residues) which is fused with a sequence derived from  
CC human osteoprotegerin (OPG; 22-194 residues) by a linker. OPG negatively  
CC regulates the formation of osteoclasts in vitro and in vivo. It blocks  
CC the differentiation of osteoclasts from monocyte or macrophage precursors  
CC and the reabsorption of bone. The OPG-Fc fusion protein is administered  
CC for the treatment of bone loss resulting from osteoporosis, Paget's  
CC disease, osteomyelitis, hypercalcaemia, osteopenia associated with  
CC surgery or steroid administration, osteonecrosis, bone loss due to  
CC rheumatoid arthritis, periodontal bone loss, osteolytic metastasis  
CC and/or prosthetic loosening.  
XX  
XX SQ Sequence 401 AA;  
  
Query Match 93.3%; Score 1245; DB 22; Length 401;  
Best Local Similarity 58.5%; Pred. No. 5.3e-35;  
Matches 233; Conservative 5; Mismatches 7; Indels 153; Gaps 2;  
  
QY 1 MDKTHCTCPAPPELLGGPSVFLFPPKPKDTLMISRTPEVTCVVDVSHEDPEVKFNWYV 60  
DB 1 MDKTHCTCPAPPELLGGPSVFLFPPKPKDTLMISRTPEVTCVVDVSHEDPEVKFNWYV 60  
  
QY 61 DGVEVHNAKTKPREEQNSTYRVWSVLTVLHODWLNKGEYCKVSNKALPAPIETISK 120  
DB 61 DGVEVHNAKTKPREEQNSTYRVWSVLTVLHODWLNKGEYCKVSNKALPAPIETISK 120  
  
QY 121 KGQPREQVYTLPPSRDELTKNQVSLTCLVKGFYPSDIAVEWESNGQPENNYKTPPVLD 180

CC cell-specific portion, comprising an humanised monoclonal antibody,  
CC having specificity for polymorphic epithelial mucin (PEM) or its antigen  
CC binding fragment and a cytotoxic portion having endonucleolytic activity,  
CC exemplified by AAM52154-AM52168 and encoded by ABA02682-ABA02728. The  
CC compound has cytostatic activity useful for treating cancer and acting as  
CC a potential inducer of apoptosis.

XX Sequence 741 AA;

Query Match 93.4%; Score 1246; DB 22; Length 741;  
Best Local Similarity 48.6%; Pred. No. 2e-34;  
Matches 233; Conservative 6; Mismatches 5; Indels 235; Gaps 3;

QY 2 DKHTCTCPAPAPLGGPSVFLPPPKDMLISRTPEVTCVVVDVSHEDPEVKENWYVD 61  
DB 241 DKHTCTCPAPAPLGGPSVFLPPPKDMLISRTPEVTCVVVDVSHEDPEVKENWYVD 300  
QY 62 GVEVHNAKTPREEQNSTYRVVSVLTVLHODWLNKGEYKCKVSNKALPAPIETISKAK 121  
DB 301 GVEVHNAKTPREEQNSTYRVVSVLTVLHODWLNKGEYKCKVSNKALPAPIETISKAK 360  
QY 122 GQREPOVYTLPPSRDELTKNQVSLTCLVKGFYPSDIAVEWESNGQPENNYKTTPPVLD 181  
DB 361 GQREPOVYTLPPSRDELTKNQVSLTCLVKGFYPSDIAVEWESNGQPENNYKTTPPVLD 420  
QY 182 DGSFFLYSKLTVDKSRWQOGNVFSCVMHEALHNHYTQKSLSLSPGK----- 228  
DB 421 DGSFFLYSKLTVDKSRWQOGNVFSCVMHEALHNHYTQKSLSLSPGKGLIAAFNIQ 480  
QY 229 ----- 228  
DB 481 TFGETKMSNATLSYIVQILSRDYDIALVQEVRSHTLVAGKLLDNLNQADPTHYVYVSE 540  
QY 229 ----- 228  
DB 541 PLGRNSYKERYLFVYRPDQVSAVDYDDGCEPCGNDTFNREPAIVRFRTFVREFA 600  
QY 229 ----- 228  
DB 601 IVPLHAAPGDAVAEIDALYDVLDOEKWGLEDMVLMGDFNACGSYVRPSQWSSIRLWTS 660  
QY 229 ---DWL-----KAFYDKV-----AEKLKEA 245  
DB 661 PTFQWLIPDSADTTATPTTHCAYDIRIVAGMLLRGAVVPDSALPFPNFAAYGLSDQLAQA 719

RESULT 10  
AAW83963  
ID AAW83963 standard; Protein; 374 AA.

AC AAW83963;

XX 05-FEB-1999 (first entry)

XX Recombinant human metFc-OB protein variant.

DE Recombinant; metFc-OB protein; Fc region; immunoglobulin; Ig: OB;  
KW obesity; human; adiposity; blood lipid; diabetes type II; insulin;  
KW hypoglycaemic; antihypertensive; diuretic; appetite suppressant;  
KW suspension; variant.

XX Homo sapiens.

XX WO9846257-A1.

XX 22-OCT-1998.

XX 16-APR-1998; 98WO-US07828.

XX 14-APR-1998; 98US-0059467.

XX 17-APR-1997; 97US-0843971.

XX (AMGE-) AMGEN INC.

XX

PI

Brems DN, French DL, Speed MA;

XX

WPI; 1998-594525/50.

DR

N-PSDB; AAV69686.

XX

Concentrated suspension of fusion of obesity protein with Fc  
immunoglobulin fragment - stable at physiological pH, used for e.g.  
reduction of weight and blood lipid levels, and for treatment of  
type II diabetes

PT

Claim 2; Fig 6A-C; 47pp; English.

XX

This represents a recombinant metFc-OB protein variant which consists of  
an Fc region of human immunoglobulin (Ig) fused to a human OB (obesity)  
protein. The invention provides a human OB protein suspension that  
contains at least 0.5 mg/ml of the human OB protein derivatised by  
attachment of the Fc region of an Ig to the N-terminus of OB, and has a  
pH 6-8. The suspensions are used to reduce weight, adiposity and blood  
lipid levels, to treat or prevent diabetes type II, and to increase lean  
mass and insulin sensitivity. They may be used in conjunction with  
insulin, hypoglycaemics, antihypertensives, diuretics, appetite  
suppressants etc. These suspensions are stable and active at  
physiological pH and are ready-for-use formulations that do not require  
freezing or freeze drying. As they are very concentrated, only small  
volumes are required and they provide a sustained-release effect, with  
increased potency and reduced frequency of injection.

XX Sequence 374 AA;

Query Match 93.3%; Score 1245; DB 19; Length 374;  
Best Local Similarity 74.4%; Pred. No. 4.5e-35;  
Matches 233; Conservative 4; Mismatches 7; Indels 69; Gaps 2;

QY 1 MDKTHCTCPAPAPLGGPSVFLPPPKDMLISRTPEVTCVVVDVSHEDPEVKFNWYV 60

DB 1 MDKTHCTCPAPAPLGGPSVFLPPPKDMLISRTPEVTCVVVDVSHEDPEVKFNWYV 60

QY 61 DGVEVINAKTPREEQNSTYRVVSVLTVLHODWLNKGEYKCKVSNKALPAPIETISKAK 120

DB 61 DGVEVINAKTPREEQNSTYRVVSVLTVLHODWLNKGEYKCKVSNKALPAPIETISKAK 120

QY 121 KGQREPOVYTLPPSRDELTKNQVSLTCLVKGFYPSDIAVEWESNGQPENNYKTTPPVLD 180

DB 121 KGQREPOVYTLPPSRDELTKNQVSLTCLVKGFYPSDIAVEWESNGQPENNYKTTPPVLD 180

QY 181 SDGSFFLYSKLTVDKSRWQOGNVFSCVMHEALHNHYTQKSLSLSPGK----- 228

DB 181 SDGSFFLYSKLTVDKSRWQOGNVFSCVMHEALHNHYTQKSLSLSPGKVPQKVQDDTKT 240

QY 229 -----DWLKAFFYDKV----- 238

DB 241 LIKTIIVTRINDISHTQSVSSKQKVTGLDPIPLGLHPIILTSKMDQTLAVYQOILTSMPSRN 300

QY 239 -----AEKLKE 244

DB 301 VIQISNDLENLRD 313

RESULT 11

AAW49075

ID AAW49075 standard; Protein; 374 AA.

XX AAW49075;

XX 18-NOV-1998 (first entry)

XX Recombinant human MetFc-OB variant 2 protein.

XX Recombinant human MetFc-OB variant 2 protein; chimeric; immunoglobulin;  
KW high blood lipid level; arterial sclerosis; stroke; Fc-OB fusion protein;  
KW diabetes.

QY 1 MDKTHCTPCPAPPELLGGPSVFLFPKPKDPTLMISRTPEVTCVVDVSHEDPEVKFNWYV 60  
 DB 1 MDKTHCTPCPAPPELLGGPSVFLFPKPKDPTLMISRTPEVTCVVDVSHEDPEVKFNWYV 60  
 QY 61 DGVEVHNAKTPREEQYNSTYRVVSVLTVLHQDLNGKEYCKVSNKALPAPIETISKAK 120  
 DB 61 DGVEVHNAKTPREEQYNSTYRVVSVLTVLHQDLNGKEYCKVSNKALPAPIETISKAK 120  
 QY 121 KGPREPOQVYTLPPSRDELTKNOVSLTCLVKGFPSPDIAVWESNGQPNENYKTTTPPVLD 180  
 DB 121 KGPREPOQVYTLPPSRDELTKNOVSLTCLVKGFPSPDIAVWESNGQPNENYKTTTPPVLD 180  
 QY 181 SDGSFFLYSKLTVDKSRWQOGNVSFSCVMHEALHNNHYTKLSLSLSPGK----- 228  
 DB 181 SDGSFFLYSKLTVDKSRWQOGNVSFSCVMHEALHNNHYTKLSLSLSPGKGGGIEGPTLR 240  
 QY 229 DWLKA 233  
 DB 241 QWLAA 245  
 RESULT 8  
 ID AAM52156 standard; Protein; 731 AA.  
 AC AAM52156;  
 XX AC  
 DT 05-FEB-2002 (first entry)  
 XX DT  
 DE Humanised HMFG-1 heavy chain/DNase I fusion protein 1.  
 XX DE  
 KW Humanised monoclonal antibody; polymorphic epithelial mucin; PEM1;  
 KW cytotoxic; endonuclease; DNase I; human; cytostatic; cancer; apoptosis.  
 XX KW  
 OS Homo sapiens.  
 OS Synthetic.  
 XX OS  
 PN WO200174905-A1.  
 XX PN  
 PD 11-OCT-2001.  
 XX PD  
 PF 26-MAR-2001; 2001WO-GB01324.  
 XX PF  
 PR 03-APR-2000; 2000GB-0008049.  
 XX PR  
 PR 02-OCT-2000; 2000US-237159P.  
 XX PR  
 XX (ANTI-) ANTISOMA RES LTD.  
 PA XX  
 PI Young RJ;  
 XX PI  
 DR WPI; 2001-662969/76.  
 XX DR  
 PT Novel compound used to treat cancer has target cell-specific portion  
 PT comprising humanised monoclonal antibody having specificity for  
 PT polymorphic epithelial mucin, and cytotoxic portion having  
 PT endonucleolytic activity  
 XX PT  
 PS Claim 20; Figure 7; 176pp; English.  
 XX PS  
 CC The invention relates to a compound which comprises a target  
 CC cell-specific portion, comprising an humanised monoclonal antibody,  
 CC having specificity for polymorphic epithelial mucin (PEM) or its antigen  
 CC binding fragment and a cytotoxic portion having endonucleolytic activity,  
 CC exemplified by AAM52154-AAM52168 and encoded by ABA02682-ABA02728. The  
 CC compound has cytostatic activity useful for treating cancer and acting as  
 CC a potential inducer of apoptosis.  
 XX CC  
 SQ Sequence 731 AA;  
 Query Match 93.4%; Score 1246; DB 22; Length 731;  
 Best Local Similarity 48.6%; Pred. No. 1.9e-34;  
 Matches 233; Conservative 6; Mismatches 5; Indels 235; Gaps 3;

QY 2 DKTHCTPCPAPPELLGGPSVFLFPKPKDPTLMISRTPEVTCVVDVSHEDPEVKFNWYV 61  
 DB 241 DKTHCTPCPAPPELLGGPSVFLFPKPKDPTLMISRTPEVTCVVDVSHEDPEVKFNWYV 300  
 QY 62 GVEVHNAKTPREEQYNSTYRVVSVLTVLHQDLNGKEYCKVSNKALPAPIETISKAK 121  
 DB 301 GVEVHNAKTPREEQYNSTYRVVSVLTVLHQDLNGKEYCKVSNKALPAPIETISKAK 360  
 QY 122 GPREPOQVYTLPPSRDELTKNOVSLTCLVKGFPSPDIAVWESNGQPNENYKTTTPPVLD 181  
 DB 361 GPREPOQVYTLPPSRDELTKNOVSLTCLVKGFPSPDIAVWESNGQPNENYKTTTPPVLD 420  
 QY 182 DGSFFLYSKLTVDKSRWQOGNVSFSCVMHEALHNNHYTKLSLSLSPGK----- 228  
 DB 421 DGSFFLYSKLTVDKSRWQOGNVSFSCVMHEALHNNHYTKLSLSLSPGKGGGLKIAAFNIQ 480  
 QY 229 ----- 228  
 DB 481 TFGETKMSNATLVSYIVQILSRDYDIALVOEVRDLSHTAVGKLLDNLNQDAPDPTTYHYVVS 540  
 QY 229 ----- 228  
 DB 541 PLGRNSYKERYLFVYRPDQVSAVDYDDGCEPCGNDTFNREPATVRFSETEVREFA 600  
 QY 229 ----- 228  
 DB 601 IVPLHAAPGDAVAIDALYDVLDVQEKWGLDVMGMDFNAGCSYVRPSQWSSIRLWTS 660  
 QY 229 ---DWL-----KAFYDKV-----AEKLUKEA 245  
 DB 661 PTQWLIPDSADTTATPTTHCAYDRIVVAGMLLRGAVVPDSALPFPNFAAYGLSDQLAAQ 719  
 RESULT 9  
 ID AAM52159 standard; Protein; 741 AA.  
 XX AC  
 XX AC  
 XX AAM52159;  
 XX DT  
 DT 05-FEB-2002 (first entry)  
 XX DT  
 DE Humanised HMFG-1 heavy chain/DNase I fusion protein 4.  
 XX DE  
 KW Humanised monoclonal antibody; polymorphic epithelial mucin; PEM1;  
 KW cytotoxic; endonuclease; DNase I; human; cytostatic; cancer; apoptosis.  
 XX KW  
 OS Homo sapiens.  
 OS Synthetic.  
 XX OS  
 PN WO200174905-A1.  
 XX PN  
 PD 11-OCT-2001.  
 XX PD  
 PF 26-MAR-2001; 2001WO-GB01324.  
 XX PF  
 PR 03-APR-2000; 2000GB-0008049.  
 XX PR  
 PR 02-OCT-2000; 2000US-237159P.  
 XX PR  
 PA (ANTI-) ANTISOMA RES LTD.  
 XX PA  
 PI Young RJ;  
 XX PI  
 DR WPI; 2001-662969/76.  
 XX DR  
 PT Novel compound used to treat cancer has target cell-specific portion  
 PT comprising humanised monoclonal antibody having specificity for  
 PT polymorphic epithelial mucin, and cytotoxic portion having  
 PT endonucleolytic activity  
 XX PT  
 PS Claim 20; Figure 10; 176pp; English.  
 XX PS  
 CC The invention relates to a compound which comprises a target

KW EPO; erythropoietin; TPO; tumour necrosis factor alpha inhibitor;  
 KW TNF-alpha inhibitor; interleukin 1 antagonist; IL-1 antagonist; TNP;  
 KW TPO mimetic peptide; EPO mimetic peptide; EMP; VEGF antagonist;  
 KW MMP inhibitor; antiinflammatory; antitumour; immunosuppressive;  
 KW cytotatic; antirheumatic; antiarthritic; antidiabetic; ophthalmological;  
 KW antianaemic; anorectic; antinfertility; haemostatic; dermatological;  
 KW neuroprotective; inflammatory disease; autoimmune disease; tumour growth;  
 KW cancer; rheumatoid arthritis; diabetic retinopathy; infertility; obesity;  
 KW sleep disorder; neurological degenerative disease; anaemia;  
 KW thrombocytopaenia; metastatic tumour; systemic lupus erythematosus;  
 KW Fanconi's syndrome.

xx Homo sapiens.  
 OS Synthetic.

xx WO200183525-A2.

xx 08-NOV-2001.

xx 02-MAY-2001; 2001WO-US14310.

xx 03-MAY-2000; 2000US-0563286.

xx (AMGE-) AMGEN INC.

xx Feige U, Liu C, Cheetham JC, Boone TC, Gudas JM;

xx WPI; 2002-130313/17.

xx N-PSDB; ABL35762.

xx Novel vehicle-peptide molecule or its multimers useful for treating  
 PT inflammatory and autoimmune diseases, cancer, rheumatoid arthritis,  
 PT diabetic retinopathy, obesity, sleep disorders and infertility

xx Example 2; Fig 8; 176pp; English.

xx The present invention describes a vehicle-peptide molecule (I) or its  
 CC multimers. (I) can have antiinflammatory, antitumour, immunosuppressive,  
 CC cytostatic, antirheumatic, antiarthritic, antidiabetic, ophthalmological,  
 CC antianaemic, anorectic, antinfertility, haemostatic, dermatological and  
 CC neuroprotective activities. (I) can be used as a therapeutic or  
 CC prophylactic agent as well as for screening purposes. (I) is useful for  
 CC diagnosing diseases characterised by dysfunction of their associated  
 CC protein of interest, for identifying normal or abnormal proteins of  
 CC interest, as a part of diagnostic kit to detect the presence of their  
 CC proteins of interest in a biological sample. Additionally, (I) is useful  
 CC for treating inflammatory and autoimmune diseases, tumour growth, cancer,  
 CC rheumatoid arthritis, diabetic retinopathy, obesity, sleep disorders,  
 CC infertility, and neurological degenerative diseases. (I), comprising  
 CC EPO-mimetic compounds are useful for treating disorders characterised by  
 CC low red blood cell levels such as anaemia. The TPO-mimetic comprising  
 CC compounds are useful for treating conditions that involve an existing  
 CC megakaryocyte/platelet deficiency or an expected megakaryocyte/platelet  
 CC deficiency, such as thrombocytopaenia, aplastic anaemia, metastatic  
 CC tumour which result in thrombocytopaenia, systemic lupus erythematosus,  
 CC and Fanconi's syndrome. ABB72403 to ABB73426 and ABL35695 to ABL35777  
 CC represent amino acid and nucleic acid sequences used in the  
 CC exemplification of the present invention.

xx Sequence 268 AA;

Query Match 93.4%; Score 1246; DB 23; Length 268;

Best Local Similarity 94.3%; Pred. No. 1.9e-35;

Matches 231; Conservative 0; Mismatches 2; Indels 12; Gaps 1;

QY 1 MDKTHCTPCPPAPPELLGGPSVFLFPPKPKDTLMISRTPEVTCVVVDVSHEDPEVKENWYV 60

DB 1 MDKTHCTPCPPAPPELLGGPSVFLFPPKPKDTLMISRTPEVTCVVVDVSHEDPEVKENWYV 60

QY 61 DGVEVHNAKTKPREQYNSTYRVVSVLTVQLQDLNKGKEYCKVSNKALPAPIETKISKA 120

DB 61 DGVEVHNAKTKPREQYNSTYRVVSVLTVQLQDLNKGKEYCKVSNKALPAPIETKISKA 120

QY 121 KGQPREQVYTLPPSRDELTKNQVSLTCLVKGFYPSDIAVEWESNGQPENNYKTPPVLD 180  
 DB 121 KGQPREQVYTLPPSRDELTKNQVSLTCLVKGFYPSDIAVEWESNGQPENNYKTPPVLD 180  
 QY 181 SDGSFFLYSKLTVDKSRWQQGNVSCVMHEALHNNHYTKQSLSPGK----- 228  
 DB 181 SDGSFFLYSKLTVDKSRWQQGNVSCVMHEALHNNHYTKQSLSPGKGGGEGTPTLR 240  
 QY 229 DWLKA 233  
 DB 241 QWLAA 245

RESULT 7

AAI96531

ID AAI96531 standard; Protein; 269 AA.

XX AC AAI96531;

XX DT 04-SEP-2000 (first entry)

XX DE Human IgG1 Fc TNP fusion protein.

XX KW Immunoglobulin; IgG1; Fc; thrombopoietin; mimetic; TNP; TPO; platelet;  
 KW megakaryocyte; production; anti-human immunodeficiency virus; anti-HIV;  
 KW anti-anaemic; dermatological; immunosuppressive; anti-inflammatory.

XX OS Homo sapiens.

XX PN WO200024770-A2.

XX PD 04-MAY-2000.

XX PF 22-OCT-1999; 99WO-US24834.

XX PR 23-OCT-1998; 98US-0105348.

XX PA (AMGE-) AMGEN INC.

XX PI Liu C, Feige U, Cheetham J;

XX WPI; 2000-365108/31.

XX N-PSDB; AAA29229.

XX Thrombopoietic peptides which activate mpl receptors and increase the  
 PT production of platelets or platelet precursors, useful for treatment of  
 PT diseases which involve thrombocytopaenia

XX Example 2A; Page 49-50; 91pp; English.

XX A compound which binds to an mpl receptor comprising a thrombopoietin  
 CC mimetic peptide (TMP) dimer joined by a linker [TMP<sub>1</sub>-1-(L<sub>1</sub>)-TMP<sub>2</sub>],  
 CC is new. TMP<sub>1</sub> and TMP<sub>2</sub> are amino acid sequences varying from at least  
 CC 10 to 14 residues in length comprising X<sub>2</sub>-X<sub>1</sub>-O, X<sub>2</sub>-X<sub>1</sub>-1, X<sub>2</sub>-X<sub>1</sub>-2,  
 CC X<sub>2</sub>-X<sub>1</sub>-3, X<sub>2</sub>-X<sub>1</sub>-4, X<sub>1</sub>-X<sub>1</sub>-O, X<sub>1</sub>-X<sub>1</sub>-1, X<sub>1</sub>-X<sub>1</sub>-2, X<sub>1</sub>-X<sub>1</sub>-3, and  
 CC X<sub>1</sub>-X<sub>1</sub>-4. X<sub>1</sub> = I, A, V, L, S or R; X<sub>2</sub> = E, K or V; X<sub>3</sub> = G or A;  
 CC X<sub>4</sub> = P; X<sub>5</sub> = T or S; X<sub>6</sub> = L, I, V, A or F; X<sub>7</sub> = R or K; X<sub>8</sub> = Q, N,  
 CC or E; X<sub>9</sub> = W, Y or F; X<sub>10</sub> = L, I, V, A, F, M or K; X<sub>11</sub> = A, I, V,  
 CC L, F, S, T, K, H, or E; X<sub>12</sub> = A, I, V, L, F, G, S, or Q; X<sub>13</sub> = R, K,  
 CC T, V, N, Q or G; X<sub>14</sub> = A, I, V, L, F, T, R, E, or G; L<sub>1</sub> = linker  
 CC comprising 1 to 20 amino acids; and n = 0 or 1. The compounds bind to and  
 CC activate the c-Mpl receptor which mediates the activity of endogenous  
 CC thrombopoietin. The TMPs are useful for increasing the production of  
 CC platelets or platelet precursors (e.g. megakaryocytes) in a mammal, which  
 CC is useful for treatment of diseases which involve thrombocytopaenia, e.g.  
 CC aplastic anaemia, immune thrombocytopaenia (ITP), human immunodeficiency  
 CC virus associated ITP, and systemic lupus erythematosus.

XX Sequence 269 AA;

Query Match 93.4%; Score 1246; DB 21; Length 269;

Best Local Similarity 94.3%; Pred. No. 1.9e-35;

Matches 231; Conservative 0; Mismatches 2; Indels 12; Gaps 1;



CC neuroprotective activities. (I) can be used as a therapeutic or  
CC prophylactic agent as well as for screening purposes. (I) is useful for  
CC diagnosing diseases characterised by dysfunction of their associated  
CC protein of interest, for identifying normal or abnormal proteins of  
CC interest, as a part of diagnostic kit to detect the presence of their  
CC proteins of interest in a biological sample. Additionally, (I) is useful  
CC for treating inflammatory and autoimmune diseases, tumour growth, cancer,  
CC rheumatoid arthritis, diabetic retinopathy, obesity, sleep disorders,  
CC infertility, and neurological degenerative diseases. (I), comprising  
CC EPO-mimetic compounds are useful for treating disorders characterised by  
CC low red blood cell levels such as anaemia. The EPO-mimetic comprising  
CC compounds are useful for treating conditions that involve an existing  
CC megakaryocyte/platelet deficiency or an expected megakaryocyte/platelet  
CC deficiency, such as thrombocytopenia, aplastic anaemia, metastatic  
CC tumour which result in thrombocytopenia, systemic lupus erythematosus,  
CC and Fanconi's syndrome. ABB72403 to ABB73426 and ABL35695 to ABL35777  
CC represent amino acid and nucleic acid sequences used in the  
CC exemplification of the present invention.

XX Sequence 248 AA;

Query Match 93.4%; Score 1246; DB 23; Length 248;  
Best Local Similarity 93.9%; Pred. No. 1.6e-35;  
Matches 230; Conservative 3; Mismatches 6; Indels 6; Gaps 1;

QY 1 MDKTHTCPCPAPPELLGGPSVFLFPPPKPDKDTLMISRTPEVTCVVDVSHEDPEVKFNWYV 60  
DB 1 MDKTHTCPCPAPPELLGGPSVFLFPPPKPDKDTLMISRTPEVTCVVDVSHEDPEVKFNWYV 60  
QY 61 DGVEVHNATKPREQYNSTYRVSVLTVLHODWLNKGEYCKVSNKALPAPIEKTIK 120  
DB 61 DGVEVHNATKPREQYNSTYRVSVLTVLHODWLNKGEYCKVSNKALPAPIEKTIK 120  
QY 121 KGQPREQVYTLPPSRDELTKNOVSLTCLVKGFYPSDIAVEWESNGQPENNYKTTTPVLD 180  
DB 121 KGQPREQVYTLPPSRDELTKNOVSLTCLVKGFYPSDIAVEWESNGQPENNYKTTTPVLD 180  
QY 181 SGGSFFLYSKLTVDKSRWQGNVFCVSMHEALHNHYTKLSLSPGK-----DWLKAF 234  
DB 181 SGGSFFLYSKLTVDKSRWQGNVFCVSMHEALHNHYTKLSLSPGKGGGFWEFTPGY 240  
QY 235 YDKVA 239  
DB 241 WQPYA 245

RESULT 5  
AAB16959  
ID AAB16959 standard; Protein; 268 AA.  
AC AAB16959;  
XX  
XX 31-OCT-2000 (first entry)  
XX  
XX FC-TMP-TMP protein sequence SEQ ID NO:8.

XX Modified peptide; therapeutic agent; fusion; Fc domain; cancer;  
KW autoimmune disease; cytostatic; antiasthmatic; thrombolytic; VEGF;  
KW immunosuppressive; EPO; TPO; CTLA4; mimetic; IL-1; TNF; antagonist;  
KW MMP; inhibitor; erythropoietin; thrombopoietin; interleukin 1;  
KW cytotoxic T cell lymphocyte antigen 4; tumour necrosis factor;  
KW vascular endothelial growth factor; matrix metalloproteinase;  
KW asthma; thrombosis; pharmaceutical.

XX Homo sapiens.  
OS Synthetic.  
OS  
XX WO200024782-A2.  
XX  
XX 04-MAY-2000.  
XX  
XX 25-OCT-1999; 99WO-US25044.

PR 23-OCT-1998; 98US-0105371.  
PR 22-OCT-1999; 99US-0428082.  
XX (AMGE-) AMGEN INC.  
XX Feige U, Liu C, Cheetham J, Boone TC;  
PI WPI; 2000-350702/30.  
XX DR N-PSDB; AAB69445.  
XX Novel composition of matter comprising an Fc domain and  
PT pharmacologically active peptides, useful for treating cancer and  
PT autoimmune diseases -  
XX Example 2; Page 182-183; 608pp; English.  
XX The present invention describes composition of matter (I) comprising an  
CC Fc domain, pharmacologically active peptides, and linkers. Where (I) is:  
CC (X1)a-F1-(X2)b, where: F1 = an Fc domain; X1 and X2 = are each  
CC independently selected from -(L1)c-P1, -(L1)c-P1-(L2)d-P2,  
CC -(L1)c-P1-(L2)d-P2-(L3)e-P3, or -(L1)c-P1-(L2)d-P2-(L3)e-P3-(L4)f-P4  
CC where P1, P2, P3, and P4 = are each independently sequences of  
CC pharmacologically active peptides; L1, L2, L3, and L4 = are each  
CC independently linkers; and a, b, c, d, e, and f = are each independently  
CC 0 or 1, provided that at least 1 of a and b is 1. The composition can  
CC have cytostatic, antiasthmatic, thrombolytic and immunosuppressive  
CC activities. DNAs, vectors and host cells from the present invention can  
CC be used for producing pharmaceutical compositions. The compositions are  
CC useful for treating cancer, asthma, thrombosis, or autoimmune diseases.  
CC The use of an Fc domain (rather than a Fab domain) can provide a longer  
CC half-life or incorporate functions such as Fc receptor binding, protein  
CC A binding, complement fixation, and possibly placental transfer. AAB69443  
CC to AAB69526 and AAB16955 to AAB18003 represent nucleotide and amino acid  
CC sequences used in the exemplification of the present invention.

XX Sequence 268 AA;

Query Match 93.4%; Score 1246; DB 21; Length 268;  
Best Local Similarity 94.3%; Pred. No. 1.9e-35;  
Matches 231; Conservative 0; Mismatches 2; Indels 12; Gaps 1;

QY 1 MDKTHTCPCPAPPELLGGPSVFLFPPPKPDKDTLMISRTPEVTCVVDVSHEDPEVKFNWYV 60  
DB 1 MDKTHTCPCPAPPELLGGPSVFLFPPPKPDKDTLMISRTPEVTCVVDVSHEDPEVKFNWYV 60  
QY 61 DGVEVHNATKPREQYNSTYRVSVLTVLHODWLNKGEYCKVSNKALPAPIEKTIK 120  
DB 61 DGVEVHNATKPREQYNSTYRVSVLTVLHODWLNKGEYCKVSNKALPAPIEKTIK 120  
QY 121 KGQPREQVYTLPPSRDELTKNOVSLTCLVKGFYPSDIAVEWESNGQPENNYKTTTPVLD 180  
DB 121 KGQPREQVYTLPPSRDELTKNOVSLTCLVKGFYPSDIAVEWESNGQPENNYKTTTPVLD 180  
QY 181 SGGSFFLYSKLTVDKSRWQGNVFCVSMHEALHNHYTKLSLSPGK-----228  
DB 181 SGGSFFLYSKLTVDKSRWQGNVFCVSMHEALHNHYTKLSLSPGKGGGGIEGPTLR 240  
QY 229 DWLKA 233  
DB 241 QWLAA 245

RESULT 6  
ABB73412  
ID ABB73412 standard; Protein; 268 AA.  
XX ABB73412;  
XX  
XX 05-APR-2002 (first entry)  
XX FC-TMP-TMP amino acid SEQ ID NO:8.  
XX Modified peptide; mimetic; Fc domain; fusion; Immunoglobulin G; IgG;  
KW

Db 181 SDGSFFLYSKLTVDKSRWQGNVFCSCVMHEALHNHYTKLSLSPOKGGGGEGPTLR 240  
 QY 229 DWLKA 233  
 Db 241 QWLAA 245

RESULT 3  
 ID AAB17953  
 AC AAB17953 standard; Protein; 248 AA.  
 XX AAB17953;  
 XX 31-OCT-2000 (first entry)  
 DE FC-IL-1 antagonist fusion protein sequence SEQ ID NO:1060.  
 XX Modified peptide; therapeutic agent; fusion; Fc domain; cancer;  
 KW autoimmune disease; cytostatic; antiasthmatic; thrombolytic; VEGF;  
 KW immunosuppressive; EPO; TPO; CTLA4; mimetic; IL-1; TNF; antagonist;  
 KW MMP; inhibitor; erythropoietin; thrombopoietin; interleukin 1;  
 KW cytotoxic T cell lymphocyte antigen 4; tumour necrosis factor;  
 KW vascular endothelial growth factor; matrix metalloproteinase;  
 KW asthma; thrombosis; pharmaceutical.  
 XX Synthetic.  
 OS  
 XX WO200024782-A2.  
 PN  
 XX 04-MAY-2000.  
 PD  
 XX 25-OCT-1999; 99WO-US25044.  
 PF  
 XX 23-OCT-1998; 98US-0105371.  
 PR  
 XX 22-OCT-1999; 99US-0428082.  
 PA (AMGE-) AMGEN INC.  
 XX Feige U, Liu C, Cheetham J, Boone TC;  
 PI WPI; 2000-350702/30.  
 XX N-PSDB; AAA69503.  
 DR Novel composition of matter comprising an Fc domain and  
 PT pharmacologically active peptides, useful for treating cancer and  
 PT autoimmune diseases -  
 XX  
 XX Example 5; Page 574-575; 608pp; English.  
 CC The present invention describes composition of matter (I) comprising an  
 CC Fc domain, pharmacologically active peptides, and linkers. Where (I) is:  
 CC (X1)-a-F1-(X2)-b, where: F1 = an Fc domain; X1 and X2 = are each  
 CC independently selected from -(L1)-c-P1, -(L1)-c-P1-(L2)-d-P2,  
 CC -(L1)-c-P1-(L2)-d-P2-(L3)-e-P3, or -(L1)-c-P1-(L2)-d-P2-(L3)-e-P3-(L4)-f-P4  
 CC where P1, P2, P3, and P4 = are each independently sequences of  
 CC pharmacologically active peptides; L1, L2, L3, and L4 = are each  
 CC independently linkers; and a, b, c, d, e, and f = are each independently  
 CC 0 or 1, provided that at least 1 of a and b is 1. The composition can  
 CC have cytostatic, antiasthmatic, thrombolytic and immunosuppressive  
 CC activities. DNAs, vectors and host cells from the present invention can  
 CC be used for producing pharmaceutical compositions. The compositions are  
 CC useful for treating cancer, asthma, thrombosis, or autoimmune diseases.  
 CC The use of an Fc domain (rather than a Fab domain) can provide a longer  
 CC half-life or incorporate functions such as Fc receptor binding, protein  
 CC A binding, complement fixation, and possibly placental transfer. AAA69443  
 CC to AAA69526 and AAB16955 to AAB18003 represent nucleotide and amino acid  
 CC sequences used in the exemplification of the present invention.  
 XX  
 XX Sequence 248 AA;  
 SQ

Query Match 93.4%; Score 1246; DB 21; Length 248;  
 Best Local Similarity 93.9%; Pred. No. 1.6e-35;  
 Matches 230; Conservative 3; Mismatches 6; Indels 6; Gaps 1;

QY 1 MDKTHTCPPCPAPPELLGGPSVFLFPPPKPDTLMISRTPEVTVVVDVSHEDPEVKFNWYV 60  
 Db 1 MDKTHTCPPCPAPPELLGGPSVFLFPPPKPDTLMISRTPEVTVVVDVSHEDPEVKFNWYV 60  
 QY 61 DGVEVHNAKTKPREEQYNSTYRVVSVLTVLRQDWLNGKEYKCKVSNKALPAPIEKTISKA 120  
 Db 61 DGVEVHNAKTKPREEQYNSTYRVVSVLTVLRQDWLNGKEYKCKVSNKALPAPIEKTISKA 120  
 QY 121 KGQPREPOVYTLPPSRDELTKNQVSLTCLVKGFYPSDIAVEWESNGQPENNYKTTPPVLD 180  
 Db 121 KGQPREPOVYTLPPSRDELTKNQVSLTCLVKGFYPSDIAVEWESNGQPENNYKTTPPVLD 180  
 QY 181 SDGSFFLYSKLTVDKSRWQGNVFCSCVMHEALHNHYTKLSLSPOKGGGGEGFETPPGY 240  
 Db 181 SDGSFFLYSKLTVDKSRWQGNVFCSCVMHEALHNHYTKLSLSPOKGGGGEGFETPPGY 240  
 QY 235 YDKVA 239  
 Db 241 WQPYA 245

RESULT 4  
 ID ABB73421  
 AC ABB73421 standard; Protein; 248 AA.  
 XX ABB73421;  
 XX 05-APR-2002 (first entry)  
 DT FC-interleukin 1 (IL-1) antagonist fusion nucleic acid SEQ ID NO:1059.  
 DE  
 XX Modified peptide; mimetic; Fc domain; fusion; immunoglobulin G; IgG;  
 KW EPO; erythropoietin; TPO; tumour necrosis factor alpha inhibitor;  
 KW TNF-alpha inhibitor; interleukin 1 antagonist; IL-1 antagonist; TMP;  
 KW TPO mimetic peptide; EPO mimetic peptide; EMP; VEGF antagonist;  
 KW MMP inhibitor; antiinflammatory; antitumour; immunosuppressive;  
 KW cytostatic; antirheumatic; antiarthritic; antidiabetic; ophthalmological;  
 KW antianaemic; anorectic; antinfertility; haemostatic; dermatological;  
 KW neuroprotective; inflammatory disease; autoimmune disease; tumour growth;  
 KW cancer; rheumatoid arthritis; diabetic retinopathy; infertility; obesity;  
 KW sleep disorder; neurological degenerative disease; anaemia;  
 KW thrombocytopaenia; metastatic tumour; systemic lupus erythematosus;  
 KW Fanconi's syndrome.  
 XX Homo sapiens.  
 OS Synthetic.  
 XX WO200183525-A2.  
 PN  
 XX 08-NOV-2001.  
 PD  
 XX 02-MAY-2001; 2001WO-US14310.  
 PF  
 XX 03-MAY-2000; 2000US-0563286.  
 PR (AMGE-) AMGEN INC.  
 XX Feige U, Liu C, Cheetham JC, Boone TC, Gudas JM;  
 PI WPI; 2002-130313/17.  
 XX N-PSDB; ABL35771.  
 DR Novel vehicle-peptide molecule or its multimers useful for treating  
 XX inflammatory and autoimmune diseases, cancer, rheumatoid arthritis,  
 PT diabetic retinopathy, obesity, sleep disorders and infertility -  
 PT  
 XX Example 5; Fig 21A-B; 176pp; English.  
 PS  
 XX The present invention describes a vehicle-peptide molecule (I) or its  
 CC multimers. (I) can have antiinflammatory, antitumour, immunosuppressive,  
 CC cytostatic, antirheumatic, antiarthritic, antidiabetic, ophthalmological,  
 CC antianaemic, anorectic, antinfertility, haemostatic, dermatological and